

```
Query Match          40.8%; Score 226; DB 15; Length 562;
Best Local Similarity 47.7%; Pred. No. 9e-18; 36; Indels 0; Gaps 0;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNR 61
DB 126 TCYEQG:SYRGTSWSTAESGAECTWNSSALAKPYSGRRPDAIRLGLGNHNYCRNPDNR 185
QY 62 RPKCYVQVGLKPLVQECMVHDCADG 87
DB 186 SKPCYVFKAGKYSSEFCSTPACSEG 211

RESULT 28
US-10-057-951-3
; Sequence 3, Application US/10057951
; Publication No. US20020177213A1
; GENERAL INFORMATION:
; APPLICANT: Moore et al.
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: PF378P1
; CURRENT APPLICATION NUMBER: US/10/057,951
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/411,977
; PRIOR FILING DATE: 1999-10-04
; PRIOR APPLICATION NUMBER: US 09/084,491
; PRIOR FILING DATE: 1998-05-27
; PRIOR APPLICATION NUMBER: US 60/048,000
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 3
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-951-3

Query Match          39.4%; Score 218.5; DB 14; Length 326;
Best Local Similarity 44.7%; Pred. No. 3.7e-17;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;

QY 3 CYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNR 62
DB 25 CYFGNGSAYRGTHSLTESGASCLPWSMILIGKVYTAQNPSAQLGLGKHNYCRNPDGDA 84
QY 63 RPKCYVQVGLKPLVQECMVHDCAD-GKKPSSPPE 95
DB 85 KPWCHLVKNRRLTWECYCDVPSCTCGLRQYSQPQ 118

RESULT 29
US-09-987-457-10
; Sequence 10, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Wether, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 10
; LENGTH: 354
; TYPE: PRT

US-09-987-457-10
; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: part of the
; OTHER INFORMATION: recombinant K2S molecule
US-09-987-457-10

Query Match          39.4%; Score 218.5; DB 11; Length 354;
Best Local Similarity 44.7%; Pred. No. 4e-17;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;

QY 3 CYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNR 62
DB 7 CYFGNGSAYRGTHSLTESGASCLPWSMILIGKVYTAQNPSAQLGLGKHNYCRNPDGDA 66
QY 63 RPKCYVQVGLKPLVQECMVHDCAD-GKKPSSPPE 95
DB 67 KPWCHLVKNRRLTWECYCDVPSCTCGLRQYSQPQ 100

RESULT 30
US-09-987-455-11
; Sequence 11, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Giradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: K2S 174-527
; OTHER INFORMATION: Description of Artificial Sequence: K2S 174-527
US-09-987-455-11

Query Match          39.4%; Score 218.5; DB 11; Length 354;
Best Local Similarity 44.7%; Pred. No. 4e-17;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;

QY 3 CYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNR 62
DB 7 CYFGNGSAYRGTHSLTESGASCLPWSMILIGKVYTAQNPSAQLGLGKHNYCRNPDGDA 66
QY 63 RPKCYVQVGLKPLVQECMVHDCAD-GKKPSSPPE 95
DB 67 KPWCHLVKNRRLTWECYCDVPSCTCGLRQYSQPQ 100

Search completed: December 3, 2003, 15:05:47
Job time : 10.8558 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:02 ; Search time 63.482 Seconds
(without alignments)
1007.637 Million cell updates/sec

Title: US-09-880-503-6

Perfect score: 2257

Sequence: 1 SNELHQPVSNCDCNLGGTCV.....VSHFLPWIRSHKRENGIAL 403

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03.*

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2257	100.0	403	23	AAE16547 Human urokinase-ty
2	2243	99.4	411	6	AAAP50871 Sequence encoded b
3	2243	99.4	411	11	AAAR06244 Urokinase precursor
4	2243	99.4	411	16	AAAG2991 Pro-urokinase. Ho
5	2243	99.4	411	21	AAAY2836 Urokinase plasmino
6	2243	99.4	411	23	AAE16544 Human urokinase-ty
7	2243	99.4	431	6	AAAP50114 Sequence encoded b
8	2243	99.4	431	7	AAAP60783 Human urokinase.
9	2243	99.4	431	8	AAAP70258 Sequence of human

10	2243	99.4	431	9	AAAP80430 Deduced AA sequenc
11	2243	99.4	431	9	AAAP81204 Pro-urokinase with
12	2243	99.4	431	10	AAAP92119 Natural human prou
13	2243	99.4	431	11	AAAR04253 Human pro-urokinas
14	2243	99.4	431	11	AAAR07112 Human pro-urokinas
15	2243	99.4	431	15	AAAR47903 Pro-urokinase deri
16	2243	99.4	431	15	AAAR63141 Full length human
17	2243	99.4	431	21	AAAY99591 Human plasminogen
18	2243	99.4	431	21	AAAY50869 Human urokinase pr
19	2243	99.4	431	22	AAAB4605 Amino acid sequenc
20	2243	99.4	431	23	AAAG79460 sc-uPA. Homo sapi
21	2243	99.4	431	23	AAAU99228 Human plasminogen
22	2243	99.4	431	23	AAE17128 Human uPA protein.
23	2243	99.4	431	24	ABU56547 Lung cancer-associ
24	2243	99.4	431	24	ABU56708 Lung cancer-associ
25	2243	99.4	431	24	ABU11076 Human urokinase pl
26	2243	99.4	431	13	AAAR20537 Amidated deriv. of
27	2243	99.4	431	13	AAAR20538 Amidated deriv. of
28	2243	99.4	436	13	AAAR20536 Inhibitor resistan
29	2240	99.2	430	18	AAW24578 Sequence of prouro
30	2240	99.2	431	10	AAAP91886 Non-glycosylated p
31	2240	99.2	431	10	AAAP94764 Mutant human prour
32	2239	99.2	411	14	AAAR34584 Human plasminogen
33	2239	99.2	431	23	AAU99230 Sequence encoded b
34	2238	99.2	411	10	AAAP96146 Pro-urokinase. Ho
35	2238	99.2	411	17	AAAR92926 Human plasminogen
36	2238	99.2	411	22	AAU99229 Human plasminogen
37	2238	99.2	431	23	AAU99229 Pro-urokinase muta
38	2237	99.1	411	16	AAAR62992 Pro-urokinase muta
39	2237	99.1	411	16	AAAR62993 Pro-urokinase muta
40	2237	99.1	411	16	AAAR62999 Pro-urokinase muta
41	2237	99.1	411	16	AAAR63000 Pro-urokinase muta
42	2235	99.0	411	16	AAAR63000 Human pro-urokinas
43	2235	99.0	411	20	AAAY39343 Human pro-urokinas
44	2235	99.0	411	20	AAAY42284 Human pro-urokinas
45	2235	99.0	411	22	AAAB20489 Human pro-urokinas

ALIGNMENTS

RESULT 1

AAE16547
ID AAE16547 standard; Protein; 403 AA.

XX AAE16547;

XX 09-APR-2002 (first entry)

XX Human urokinase-type plasminogen activator scUPA delta136-143 mutant.

XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
XX stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
XX microvascular occlusion; angiotensin disorder; pulmonary fibrosis; asthma;
XX tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
XX clotting disorder; uterine contraction disorder; respiratory disease;
XX male impotence; adult respiratory distress syndrome; scUPA delta136-143;
XX single chain urokinase; mutant; mutein.

XX Homo sapiens.

OS Synthetic.

XX WO200197752-A2.

XX 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US18976.

XX 20-JUN-2000; 2000US-212874P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Cines DB, Higazi AA;

XX WPI; 2002-122240/16.
 DR N-PSDB; AAD27080.
 XX
 XX Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator -
 XX
 XX Claim 22; Fig 1F; 117pp; English.
 XX
 XX The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 CC diabetic retinopathy, wound healing, clotting disorder, uterine
 CC contraction disorder, male impotence, respiratory disease or condition
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) single
 CC chain urokinase (scuPA) deletion mutant designated as scuPA delta136-143.
 XX
 XX Sequence 403 AA;

Query Match 100.0%; Score 2257; DB 23; Length 403;
 Best Local Similarity 100.0%; Pred. No. 8.1e-176;
 Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 60
 Db 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 60
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
 Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
 QY 121 PLVQECVHDCADGKLFQCGQKTLRPRFKIIGGFTTIENOPFAAIYRRHGGSVTYV 180
 Db 121 PLVQECVHDCADGKLFQCGQKTLRPRFKIIGGFTTIENOPFAAIYRRHGGSVTYV 180
 QY 181 CGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFVENLILHKDYSAD 240
 Db 181 CGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFVENLILHKDYSAD 240
 QY 241 TLAHNDIALKIRSGRCAOPSRITQITCLPSMYNDPQFGTSCETITGFGKNSDYL 300
 Db 241 TLAHNDIALKIRSGRCAOPSRITQITCLPSMYNDPQFGTSCETITGFGKNSDYL 300
 QY 301 PEQLKXVTVVKLISHRECCQPHYVGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRM 360
 Db 301 PEQLKXVTVVKLISHRECCQPHYVGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRM 360
 QY 361 TLTGIVSWRGKALDKPGVYTRVSHFLPWIRSHRKEENGLAL 403
 Db 361 TLTGIVSWRGKALDKPGVYTRVSHFLPWIRSHRKEENGLAL 403

RESULT 2

AAFP50871
 ID AAP50871 standard; protein; 411 AA.
 XX
 AC AAP50871;
 XX
 DT 30-NOV-1991 (first entry)
 XX
 DE Sequence encoded by cDNA sequence for human urokinase zymogen
 DE (Japanese Patent Application No.37119/84).
 XX
 XX Thrombolytic agent; plasminogen activator activity; fibrin affinity;

XX enzyme.
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Cleavage-site 158..159 /note= "potential cleavage site which generates
 FT the two-chain form from the zymogen"
 FT
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..279
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380
 XX EPI39447-A.
 XX 02-MAY-1985.
 XX 07-SEP-1984; 84EP-0306117.
 XX 17-OCT-1983; 83JP-0195051.
 PR 13-SEP-1983; 83JP-0170354.
 XX (GREG) GREEN CROSS CORP.
 XX Kasai S, Arimura H, Mori K, Nishida M, Suyama T;
 PI WPI; 1985-106530/18.
 XX New urokinase zymogen - useful as thrombolytic agent
 PT Disclosure; Page 12; 30pp; English.
 PS
 CC Zymogen AAP50871 is the inactive precursor form of human urokinase.
 CC Urokinase zymogen is cleaved into the two-chain form composed of
 CC characteristic urokinase H (molecular wt. of 30,000) and L (molecular
 CC wt. of 20,000) chains when treated with catalytic amounts of plasmin.
 CC The patentors claim a new urokinase zymogen which has mol. wt. ca.
 CC 50,000, a single chain molecular structure, and selective affinity
 CC for fibrin. It is a thrombolytic agent which manifests its
 CC plasminogen activator activity on cleavage by proteolytic enzymes
 CC (e.g. plasmin) and has higher affinity for fibrin than known forms
 CC of urokinase.
 XX
 SQ Sequence 411 AA;
 Query Match 99.4%; Score 2243; DB 6; Length 411;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 60
 Db 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 60
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
 Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
 QY 121 PLVQECVHDCADGKLFQCGQKTLRPRFKIIGGFTTIENOPFAAIYRRH 172
 Db 121 PLVQECVHDCADGKLFQCGQKTLRPRFKIIGGFTTIENOPFAAIYRRH 180
 QY 173 RGGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFVENLI 232
 Db 181 RGGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFVENLI 240
 QY 233 LHKDYSADTLAHNDIALKIRSGRCAOPSRITQITCLPSMYNDPQFGTSCETITGFGK 292
 Db 241 LHKDYSADTLAHNDIALKIRSGRCAOPSRITQITCLPSMYNDPQFGTSCETITGFGK 300

```
QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTTMLCAADPQWKTDSCQDSGGPL 352
DB 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTTMLCAADPQWKTDSCQDSGGPL 360
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRVSHFLPWIRSHTKENGLAL 403
DB 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRVSHFLPWIRSHTKENGLAL 411

RESULT 3
AAR06244
ID AAR06244 standard; protein; 411 AA.
XX
AC AAR06244;
XX
DT 07-DEC-1990 (first entry)
XX
DE Urokinase precursor protein.
XX
KW Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis;
KW myocardial infarction;
XX
OS Homo sapiens.
XX
XX EP380334-A.
XX
XX 01-AUG-1990.
XX
XX 25-JAN-1990; 90EP-0300772.
XX
XX 17-MAY-1989; 89JP-0121405.
XX
XX 27-JAN-1989; 89JP-0016406.
XX
XX (GREC ) GREEN CROSS CORP.
XX
XX Matsuda K, Ueda Y, Tamanouchi K;
XX WPI; 1990-233117/31.
XX
XX Urokinase precursor-lipid composite - used as thrombolytic agent,
PT having prolonged half-life in the blood, enhanced
PT bio-availability and improved activity
XX
XX Claim 3; Fig 1; 11pp; English.
XX
XX By forming a precursor-lipid composite, the half-life of this
CC thrombolytic agent in the blood may be increased, exhibiting
CC improved activity without abnormal acceleration of fibrinolytic
CC activity. Compound is useful as a thrombolytic agent in
CC treatment of cerebral thrombosis, myocardial infarction etc.
XX
XX Sequence 411 AA;
XX
Query Match 99.4%; Score 2243; DB 11; Length 411;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGTCSNXYFNIHWCNCPKXFGGQHCIDSKTCYEGNGHYRG 60
DB 1 SNELHQVPSNCDCLNGTCSNXYFNIHWCNCPKXFGGQHCIDSKTCYEGNGHYRG 60
QY 61 KASDTMGRPCLPNWSATVLQQTVAHRSALQGLGKHNCRPNPNRRRPWCYVQVGLK 120
DB 61 KASDTMGRPCLPNWSATVLQQTVAHRSALQGLGKHNCRPNPNRRRPWCYVQVGLK 120
QY 121 PLVQECMWHCADGK-----LKFCQGGKTLRPRFKIIGGEFTTIENQPFAAIYRRH 172
DB 121 PLVQECMWHCADGKXPSSPEELKFQCGGKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
QY 173 RGSVTVVCGSLSPCWVISATHCFIDYPKKEDIYVLGRSLNSNTQGMKFEVENLI 232
DB 181 RGSVTVVCGSLSPCWVISATHCFIDYPKKEDIYVLGRSLNSNTQGMKFEVENLI 240
```

```
QY 233 LHKDYSADTLAHNDIALLKIRSKEGRCAPSRITQITCLPSMYNDPOFGTSCEITGFGK 292
DB 241 LHKDYSADTLAHNDIALLKIRSKEGRCAPSRITQITCLPSMYNDPOFGTSCEITGFGK 300
QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTTMLCAADPQWKTDSCQDSGGPL 352
DB 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTTMLCAADPQWKTDSCQDSGGPL 360
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRVSHFLPWIRSHTKENGLAL 403
DB 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRVSHFLPWIRSHTKENGLAL 411

RESULT 4
AAR62991
ID AAR62991 standard; protein; 411 AA.
XX
AC AAR62991;
XX
DT 25-MAR-2003 (updated)
DT 21-SEP-1995 (first entry)
XX
DE Pro-urokinase.
XX
KW Pro-urokinase; thrombolysis; fibrin clot lysis.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Domain 297..313
XX Disulfide-bond 11..19 /note= "flexible loop"
XX Disulfide-bond 13..31
XX Disulfide-bond 33..42
XX Disulfide-bond 50..131
XX Disulfide-bond 71..113
XX Disulfide-bond 102..126
XX Disulfide-bond 148..279
XX Disulfide-bond 189..205
XX Disulfide-bond 197..268
XX Disulfide-bond 293..362
XX Disulfide-bond 325..341
XX Disulfide-bond 352..380
XX
XX WO9501427-A1.
XX
XX 12-JAN-1995.
XX
XX 28-JUN-1994; 94WO-US07278.
XX
XX 02-JUL-1993; 93US-0087163.
XX
XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
XX
XX Gurewich V, Liu J;
XX WPI; 1995-060991/08.
XX
XX Pro-urokinase mutants - have thrombolytic activity but reduced
PT fibrinogenolysis activity and non-specific plasminogen activation
XX
XX Disclosure; Fig 1; 46pp; English.
XX
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
CC described in AAR62992-R6308 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the
CC treatment of thromboembolism, but have a reduced fibrinogenolysis
CC activity and non-specific plasminogen activation. The mutants can
CC therefore be used for the lysis of fibrin clots without inducing
CC systemic bleeding, as can be the case with the wild type protein.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
```


SQ	Sequence	411 AA;
XX	Query Match	99.4%; Score 2243; DB 16; Length 411;
XX	Best Local Similarity	98.1%; Pred. No. 1.2e-174;
XX	Matches	403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
Qy	1	SNELHQPNSDCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
Db	1	SNELHQPNSDCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
Qy	61	KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRPNDRRPPWCYVOVGLK 120
Db	61	KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRPNDRRPPWCYVOVGLK 120
Qy	121	PLVQECMWHDCADGK-----LKFCQCGKTLRPRFKIIGBETTIENQPFPAALYRRH 172
Db	121	PLVQECMWHDCADGK-----LKFCQCGKTLRPRFKIIGBETTIENQPFPAALYRRH 172
Qy	121	PLVQECMWHDCADGKPPPELKFQCGQKTLRPRFKIIGBETTIENQPFPAALYRRH 180
Db	121	PLVQECMWHDCADGKPPPELKFQCGQKTLRPRFKIIGBETTIENQPFPAALYRRH 180
Qy	173	RGGSVTYVCGGSLISPCWISATHCFDYPKKEDYIVVLGRSRLNSNTQGMKFEVENLI 232
Db	181	RGGSVTYVCGGSLISPCWISATHCFDYPKKEDYIVVLGRSRLNSNTQGMKFEVENLI 240
Qy	233	LHKDYSADTLAHHNDIALKIRSKRGCAQPSRTIQTICLPMSYNDPFGTSCIEITGFGK 292
Db	241	LHKDYSADTLAHHNDIALKIRSKRGCAQPSRTIQTICLPMSYNDPFGTSCIEITGFGK 300
Qy	293	ENSTDYLYPEOLKMTVVKLISHRCCQPHYVGSVTTKMLCAADPQWKTDSCQDSSGGPL 352
Db	301	ENSTDYLYPEOLKMTVVKLISHRCCQPHYVGSVTTKMLCAADPQWKTDSCQDSSGGPL 360
Qy	353	VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
Db	361	VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
XX	RESULT 5	
XX	ID	AA92836 standard; Protein; 411 AA.
XX	AC	AA92836;
XX	DT	29-AUG-2000 (first entry)
XX	DE	Urokinase plasminogen activator (uPA).
XX	KW	N-terminal; pro-uPA; urokinase plasminogen activator; receptor; uPAR; anti-cancer; anti-metastatic; anti-proliferative; anti-atherosclerotic; anti-thrombotic; anti-angiogenic; anti-inflammatory; anti-arthritis; anti-fibrotic; apoptotic; vasotropic; anti-diabetic; ophthalmological; thrombolytic.
XX	OS	Homo sapiens.
XX	FH	Location/Qualifiers
FT	Disulfide-bond	11..19
FT	Disulfide-bond	13..31
FT	Disulfide-bond	33..42
FT	Disulfide-bond	50..131
FT	Disulfide-bond	71..113
FT	Disulfide-bond	102..126
FT	Disulfide-bond	148..279
FT	Disulfide-bond	189..205
FT	Disulfide-bond	197..268
FT	Disulfide-bond	293..362
FT	Disulfide-bond	325..341
FT	Disulfide-bond	352..380
XX	WO200026353-A1.	
XX	PD	11-MAY-2000.
XX	28-OCT-1999;	99WO-US25210.

PR	29-OCT-1999;	98US-0181816.
XX	(ANGS-) ANGSTROM PHARM INC.	
XX	Mazar AP, Jones TR;	
XX	WPI; 2000-365605/31.	
XX	New cyclic peptide, useful for treatment or diagnosis of e.g. tumors and other diseases involving cell proliferation or migration, targets the urokinase plasminogen activator receptor	
XX	Disclosure; Fig 1; 93pp; English.	
XX	The present sequence shows the wild-type urokinase plasminogen activator (uPA). Cyclic peptides based on the amino acids residues 20-30 (the receptor-binding region) of uPA are claimed. These cyclic peptides target the uPA receptor (uPAR), allowing therapeutic or diagnostic agents to be delivered to uPAR-expressing cells. The cyclic peptides are used, optionally when linked to a therapeutic agent, to inhibit migration, invasion and proliferation of cells, or angiogenesis, or to induce apoptosis. Particularly they are used, in human or veterinary medicine, to treat diseases characterized by these processes, e.g. solid tumors, leukaemia or lymphoma (or their metastases); benign hyperplasia; atherosclerosis; restenosis; ischaemia; deep vein thrombosis; neovascular glaucoma; diabetic retinopathy; arthritis; fibrosis; bone fracture etc., most particularly growth, invasion and metastasis of tumors. When labeled, the cyclic peptides can be used for diagnostic detection of uPAR (a marker of metastasis) on cells, tissues etc., in vivo or in vitro, and when immobilized they are used to isolate uPAR or cells that express them. The cyclic peptides are stable, soluble in water, bind strongly to uPAR, are relatively inexpensive to produce and may be derivatized by attachment of therapeutic or diagnostic agents without significantly affecting their binding. Since they target uPAR, they should have relatively low systemic toxicity and only low doses are required.	
XX	Sequence	411 AA;
SQ	Query Match	99.4%; Score 2243; DB 21; Length 411;
	Best Local Similarity	98.1%; Pred. No. 1.2e-174;
	Matches	403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
Qy	1	SNELHQPNSDCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
Db	1	SNELHQPNSDCCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
Qy	61	KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRPNDRRPPWCYVOVGLK 120
Db	61	KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRPNDRRPPWCYVOVGLK 120
Qy	121	PLVQECMWHDCADGK-----LKFCQCGKTLRPRFKIIGBETTIENQPFPAALYRRH 172
Db	121	PLVQECMWHDCADGKPPPELKFQCGQKTLRPRFKIIGBETTIENQPFPAALYRRH 180
Qy	173	RGGSVTYVCGGSLISPCWISATHCFDYPKKEDYIVVLGRSRLNSNTQGMKFEVENLI 232
Db	181	RGGSVTYVCGGSLISPCWISATHCFDYPKKEDYIVVLGRSRLNSNTQGMKFEVENLI 240
Qy	233	LHKDYSADTLAHHNDIALKIRSKRGCAQPSRTIQTICLPMSYNDPFGTSCIEITGFGK 292
Db	241	LHKDYSADTLAHHNDIALKIRSKRGCAQPSRTIQTICLPMSYNDPFGTSCIEITGFGK 300
Qy	293	ENSTDYLYPEOLKMTVVKLISHRCCQPHYVGSVTTKMLCAADPQWKTDSCQDSSGGPL 352
Db	301	ENSTDYLYPEOLKMTVVKLISHRCCQPHYVGSVTTKMLCAADPQWKTDSCQDSSGGPL 360
Qy	353	VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
Db	361	VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
XX	RESULT 6	
XX	AAE16544	

ID AAE16544 standard; Protein; 411 AA.
 AC AAE16544;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Human urokinase-type plasminogen activator tcupA and scuPA protein.
 DE
 XX
 XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
 KW microvascular occlusion; angiotensin II; pulmonary fibrosis; asthma;
 KW tumour cell metastasis; glioma; diabetic retinopathy; wound healing;
 KW clotting disorder; uterine contraction disorder; respiratory disease;
 KW male impotence; adult respiratory distress syndrome; tcupA; scuPA;
 KW two chain urokinase; single chain urokinase.
 XX
 OS Homo sapiens.
 XX
 XX WO200197752-A2.
 XX
 XX 27-DEC-2001.
 XX
 XX 13-JUN-2001; 2001WO-US18976.
 XX
 XX 20-JUN-2000; 2000US-212874P.
 XX
 XX (UYPE-) UNIV PENNSYLVANIA.
 XX
 XX Cines DB, Higazi AA;
 PI
 XX WPI; 2002-122240/16.
 DR
 DR N-PSDB; AAD27077.
 XX
 XX Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator -
 XX
 XX Claim 9; Fig 1C; 117pp; English.
 XX
 XX The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiotensin II, angiotensin II, angiotensin II, angiotensin II,
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 CC diabetic retinopathy, wound healing, clotting disorder, uterine
 CC contraction disorder, male impotence, respiratory disease or condition
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) two
 CC chain urokinase (tcupA) and single chain urokinase (scuPA) protein.
 XX
 SQ Sequence 411 AA;
 Query Match 99.4%; Score 2243; DB 23; Length 411;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174; Indels 8; Gaps 1;
 Matches 403; Conservative 0; Mismatches 0;
 QY 1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
 DB 1 SNELHQVPSNCDLNGGTCVSNKYSFNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
 QY 61 KASDTMGRPLPWNATVLOOTVHAHRSALQGLGKHNVCNPDNRRPWCYVQGLK 120
 DB 61 KASDTMGRPLPWNATVLOOTVHAHRSALQGLGKHNVCNPDNRRPWCYVQGLK 120
 QY 121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH 172
 DB 121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAIYRRH 180

QY 173 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
 DB 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 240
 QY 233 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTTQITCLPSMYNDPQFGTSCETITGFGK 292
 DB 241 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTTQITCLPSMYNDPQFGTSCETITGFGK 300
 QY 293 ENSTDVLYPEQLKMTVVKLIISHRECOQPHYGVSEVTTKWLCAADPOWKTDSCQDGGGGL 352
 DB 301 ENSTDVLYPEQLKMTVVKLIISHRECOQPHYGVSEVTTKWLCAADPOWKTDSCQDGGGGL 360
 QY 353 VCSLQGRMTLTGIVSGRCALKDKPGVYTVRVSHFLPWIRSHTKENGGLAL 403
 DB 361 VCSLQGRMTLTGIVSGRCALKDKPGVYTVRVSHFLPWIRSHTKENGGLAL 411
 RESULT 7
 AAP50114
 ID AAP50114 standard; Protein; 431 AA.
 XX
 AC AAP50114;
 XX
 DT 27-SEP-1991 (first entry)
 XX
 DE Sequence encoded by the signal sequence and noncoding region of the
 DE pro-UK structural gene (Sequence II).
 XX
 KW Enzyme; thrombosis therapy; embolic disease;
 KW single-chain pro-urokinase.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 21..177
 FT Domain /label= A chain
 FT Domain 179..431
 FT Domain /label= B chain
 FT Peptide 1..20
 FT Peptide /label= signal peptide
 XX
 XX EPI54272-A.
 PN
 XX 11-SEP-1985.
 PD
 XX 23-FEB-1985; 8SEP-0102031.
 PF
 XX 31-JAN-1985; 85JP-0017969.
 PR
 XX 27-FEB-1984; 84JP-0037119.
 XX
 XX (GREG) GREEN CROSS CORP.
 PA
 XX Hiramatsu R, Kaneda T, Nagai M, Arimura H, Nishida M;
 PI Suyama T;
 XX
 XX WPI; 1985-224693/37.
 DR N-PSDB; AAN50138.
 DR
 XX Glycosylated single-chain pro-urokinase - prepd. by cultivating
 PT animal cells transformed by DNA prepd. from m RNA
 XX
 XX Disclosure; Page 8-10; 64pp; English.
 PS
 XX The inventors claim a method of producing single-chain
 CC pro-urokinase by using as template, mRNA obtd. from cells of an
 CC established human kidney-derived cell line. The urokinase is used to
 CC treat thrombosis and embolic diseases as well as in the treatment of
 CC diseases in combination with anticancer agents.
 CC
 XX Sequence 431 AA;
 SQ
 Query Match 99.4%; Score 2243; DB 6; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;

Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCEIDKSKTCYEGNGHFYRG 60
 DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCEIDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
 DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGK-----LKFCQCKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
 DB 141 PLVQECMVHDCADGKPKSPPEELKFCQCKTLRPRFKIIGGEFTTIENQWFAAIYRRH 200
 QY 173 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
 DB 201 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 233 LHKDYSADTLAHDNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
 DB 261 LHKDYSADTLAHDNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
 QY 293 ENSTDYLYPEQLKMTVVKLISHRECQOPHYVGSEVTTKMLCAADPQWKTDSCQDSGGPL 352
 DB 321 ENSTDYLYPEQLKMTVVKLISHRECQOPHYVGSEVTTKMLCAADPQWKTDSCQDSGGPL 380
 QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
 DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 8

AAP60783
 ID AAP60783 standard; Protein; 431 AA.

XX AC AAP60783;

DT 25-MAR-2003 (updated)
 DT 23-OCT-1991 (first entry)

XX DE Human urokinase.

XX KW E.coli; high molecular urokinase.

XX OS Homo sapiens.

PH Key Location/Qualifiers
 FT Protein 21..431

XX PN JP61181377-A.

XX PD 14-AUG-1986.

XX PF 25-JAN-1985; 85JP-0011032.

XX PR 25-JAN-1985; 85JP-0011032.

XX PA (NISC) NISSAN CHEM IND LTD.

PA (HODO) HODOGAYA CHEM IND CO LTD.

PA (SAGA) SAGAMI CHEM RES CENTRE.

PA (CENG) CENTRAL GLASS CO LTD.

PA (NIPS) NIPPON SODA CO.

PA (TOYO) TOYO SODA MFG CO LTD.

XX WIPI; 1986-254744/39.

DR N-PSDB; AAN60703.

XX PT Human urokinase gene - has N-end of aminoacid sequence coded by

FT codon used in Escherichia coli.

XX PS Disclosure; Fig 2; 19pp; Japanese.

XX CC The claimed gene product may be expressed in a transformed E.coli

CC host, for the efficient production of high molecular human urokinase.
 CC The N-terminal of the protein expressed by the transforming plasmid
 CC is replaced with a codon frequently used in E.coli.
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 431 AA;

Query Match 99.4%; Score 2243; DB 7; Length 431;
 Best Local Similarity 98.1%; Pred.No.1.2e-174;

Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCEIDKSKTCYEGNGHFYRG 60

DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKPKFGQHCEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120

DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGK-----LKFCQCKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172

DB 141 PLVQECMVHDCADGKPKSPPEELKFCQCKTLRPRFKIIGGEFTTIENQWFAAIYRRH 200

QY 173 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLI 232

DB 201 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLI 260

QY 233 LHKDYSADTLAHDNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292

DB 261 LHKDYSADTLAHDNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320

QY 293 ENSTDYLYPEQLKMTVVKLISHRECQOPHYVGSEVTTKMLCAADPQWKTDSCQDSGGPL 352

DB 321 ENSTDYLYPEQLKMTVVKLISHRECQOPHYVGSEVTTKMLCAADPQWKTDSCQDSGGPL 380

QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403

DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 9

AAP70258
 ID AAP70258 standard; Protein; 431 AA.

XX AC AAP70258;

XX DT 25-MAR-2003 (updated)

DT 19-MAY-1991 (first entry)

XX DE Sequence of human prourokinase and leader.

XX KW Cardiovascular disease treatment; fibrin affinity; thrombolytic;
 KW enzyme; protease.

XX OS Homo sapiens.

PH Key Location/Qualifiers

FT Peptide 1..20

FT Protein /label= leader

FT Protein 21..431

XX FT /label= prourokinase

XX PN EP231883-A.

XX PD 12-AUG-1987.

XX PF 29-JAN-1987; 87EP-0101209.

XX PR 31-JAN-1986; 86JP-0017734.

XX PR 30-JAN-1987; 87JP-0018626.

XX PA (SAGA) SAGAMI CHEM RES CENTRE.

XX PA (NIPS) NIPPON SODA CO.

(CENG) CENTRAL GLASS CO LTD.
 PA (TOYT) TOYO SODA MFG CO LTD.
 PA (NISC) NISSAN CHEM IND LTD.
 PA (NISC) NISSAN CHEMICAL INDS KK.
 XX
 XI Tagawa M, Wada M, Yamada M, Yokoyama M, Numao N;
 XX PI
 XX DR WPI; 1987-222882/32.
 DR N-PSDB; AAN70390.
 XX
 XX Hybrid plasminogen activator-like polypeptide - having a region
 PT for affinity to fibrin from tissue plasminogen activator and a
 PT region from prourokinase
 XX
 XX Disclosure; Fig 2(1-5); 64pp; English.
 PS
 XX The TPA portion of the claimed hybrid polypeptide (see FT) may
 CC consist of 2 kringles from N-terminal first serine to 219th glycine
 CC of human TPA, 1 krigle from 128th serine to 219th glycine of human
 CC TPA or half a krigle from 161st methionine to 219th glycine (see
 CC AAP70257). The C-terminal half of the hybrid polypeptide may contain
 CC an AA sequence from 150th glutamine to C-terminal 411th leucine of
 CC prourokinase (see AAP70258).
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 XX Sequence 431 AA;
 SQ

Query Match 99.4%; Score 2243; DB 8; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGQHCEDKSKTCYEGNGHFYRG 60
 DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGQHCEDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDNRRRRCWCVQVGLK 120
 DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDNRRRRCWCVQVGLK 140
 QY 121 PLVQECMWHDCADGK-----LKFQCGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRH 172
 DB 141 PLVQECMWHDCADGKPPPELKFQCGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRH 200
 QY 173 RGSVTVYVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKKEVENLI 232
 DB 201 RGSVTVYVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKKEVENLI 260
 QY 233 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGK 292
 DB 261 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGK 320
 QY 293 ENSTDYLYPEQLKMTVVVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDSDGGL 352
 DB 321 ENSTDYLYPEQLKMTVVVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDSDGGL 380
 QY 353 VCSLQGRMTLTGIVSGRGKALCKDKPGVYTRVSHFPLWIRSHTKKEENGLAL 403
 DB 381 VCSLQGRMTLTGIVSGRGKALCKDKPGVYTRVSHFPLWIRSHTKKEENGLAL 431

RESULT 10
 AAP80430
 ID AAP80430 standard; protein; 431 AA.
 XX
 AC AAP80430;
 XX
 XX 25-MAR-2003 (updated)
 DT 14-SEP-1990 (first entry)
 XX
 DE Deduced AA sequence of the single chain urokinase plasminogen activator
 DE (SCU-PA) cDNA insert prepared from human Hep3 cells.
 XX
 XX Single chain urokinase plasminogen activator (SCU-PA); human Hep3 cells;

glyceraldehyde-3-phosphate dehydrogenase gene; thrombosis prevention;
 thrombosis treatment.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 PH Peptide 1..20
 FT Protein 21..411
 XX
 XX EP288435-A.
 XX 26-OCT-1988.
 XX
 XX 11-APR-1988; 88EP-0810234.
 XX
 XX 15-APR-1987; 87GB-0009081.
 XX 16-JUN-1987; 87GB-0014059.
 XX 04-DEC-1987; 87IE-0003299.
 XX (CIBA) CIBA GEIGY AG.
 XX Meyhack B, Heim J, Burgi R;
 XX WPI; 1988-301440/43.
 XX N-PSDB; AAN80581.
 XX
 XX Prodn. of human single chain urokinase type plasminogen activator -
 PT by culturing yeast strain transformed with hybrid vector contg.
 PT yeast expression control sequences
 XX
 XX Example 1; Fig 2; 48pp; English.
 XX
 XX The patent is for the prodn. of human single chain urokinase-type
 CC plasminogen activator (UTPA). Mutants of scu-PA are especially those
 CC which render the protein protease resistant. Such scu-PA mutants are
 CC covalently modified at sites of proteolysis by proteases occurring in
 CC blood such as thrombin or plasmin, so that they are no longer
 CC susceptible to protease hydrolysis at these locations. The target sites
 CC include Lys135 to Lys136 (cleavage at this site generates the so-called
 CC low molecular weight form of scu-PA or LUK), Arg156 to Phe157
 CC (susceptible to thrombin attack) and Lys158 to Ile159 (cleavage at this
 CC site by plasmin generates tuc-PA). Suitable scu-PA mutants have site
 CC specific substitutions, insertions or deletions of residues at
 CC one or more of these target sites. Especially preferred are those
 CC mutants in which one amino acid residue or both amino acid residues
 CC forming the target sites are deleted or in which at least one of these
 CC amino acid residues is replaced by another amino acid residue so that
 CC the resulting mutants are resistant to proteolytic attack. The UTPA
 CC proteins exhibit the biological activity of natural human UTPA without
 CC any refolding procedure being necessary. They can be used as for known
 CC PAs in humans for the prevention or treatment of thrombosis or other
 CC conditions where it is desired to produce local fibrinolytic or
 CC proteolytic activity.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 XX
 XX Sequence 431 AA;
 SQ

Query Match 99.4%; Score 2243; DB 9; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGQHCEDKSKTCYEGNGHFYRG 60
 DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGQHCEDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDNRRRRCWCVQVGLK 120
 DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDNRRRRCWCVQVGLK 140
 QY 121 PLVQECMWHDCADGK-----LKFQCGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRH 172
 DB 141 PLVQECMWHDCADGKPPPELKFQCGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRH 200

QY 173 RGSVTVVCGSLISPCWVISATHCFIDYPKXEDIVYLGSRSLNSNTQGMKFEVENLI 232
 DB 201 RGSVTVVCGSLISPCWVISATHCFIDYPKXEDIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 233 LHKDYSADTLAHHNDIALKIRSKGRCAPQPSRTIQTICLPSMYNDPQGTSCITGFGK 292
 DB 261 LHKDYSADTLAHHNDIALKIRSKGRCAPQPSRTIQTICLPSMYNDPQGTSCITGFGK 320
 QY 293 ENSTDYLYPEOLKMTVVVKLISHRECQPHYVGVSEVTTKMLCAADPQWKTDSCQDSSGGPL 352
 DB 321 ENSTDYLYPEOLKMTVVVKLISHRECQPHYVGVSEVTTKMLCAADPQWKTDSCQDSSGGPL 380
 QY 353 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
 DB 381 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 11

AAP81204
 ID AAP81204 standard; protein; 431 AA.

AC AAP81204;

DT 25-MAR-2003 (updated)

DT 03-DEC-1990 (first entry)

XX Pro-urokinase with signal sequence.

XX pro-urokinase (pro-UK); plasminogen activator; pUK33; ss.

XX Key Location/Qualifiers

FT Peptide 1..20

FT Protein /label=signal peptide

FT 21..431

FT /label=pro-urokinase

XX EP265874-A.

XX 04-MAY-1988.

XX 23-OCT-1987; 87EP-0115600.

XX 23-OCT-1986; 86JP-0253078.

XX (GRC) GREEN CROSS CORP.

PI Amatsuji Y, Okabayashi K, Nagai M, Arimura H, Suyama T;

XX WPI; 1988-121000/18.

XX N-PSDB; AAN81558.

XX Glycosylated single-chain pro-urokinase prodn -

PT by cultivating DHFR gene-deficient CHO-K1 cells transformed

PT with a plasmid contg cDNA, SV40 promoter and DHFR gene.

XX Disclosure; Page ?; ?pp; English.

XX The Arg at position 2 is encoded by TGA(sic). Possible error in the

CC specification. Should read CGA ?

CC The pro-UK gene was derived from plasmid pUK33. The cDNA was

CC synthesised using urokinase mRNA isolated from a human kidney cell

CC line. Pro-UK was cloned into a SV40 promoter-contg. plasmid, down-

CC stream of the promoter. This plasmid was then ligated to a DHFR-

CC gene contg. plasmid so that pro-UK and DHFR are inserted in

CC opposite directions. The recombinant plasmid was used to transform

CC CHO-K1 cell derived DHFR gene-deficient host cells to produce

CC glycosylated single-chain pro-UK.

CC (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 431 AA;

SQ

Query Match

Best Local Similarity 99.4%; Score 2243; DB 9; Length 431;

Pred. No. 1.2e-174;

Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
 DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 80
 QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVGLK 120
 DB 81 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEBTTTIENQFWFAAIYRRH 172
 DB 141 PLVQECMVHDCADGKSPPEELKFCGQKTLRPRFKIIGGEBTTTIENQFWFAAIYRRH 200
 QY 173 RGSVTVVCGSLISPCWVISATHCFIDYPKXEDIVYLGSRSLNSNTQGMKFEVENLI 232
 DB 201 RGSVTVVCGSLISPCWVISATHCFIDYPKXEDIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 233 LHKDYSADTLAHHNDIALKIRSKGRCAPQPSRTIQTICLPSMYNDPQGTSCITGFGK 292
 DB 261 LHKDYSADTLAHHNDIALKIRSKGRCAPQPSRTIQTICLPSMYNDPQGTSCITGFGK 320
 QY 293 ENSTDYLYPEOLKMTVVVKLISHRECQPHYVGVSEVTTKMLCAADPQWKTDSCQDSSGGPL 352
 DB 321 ENSTDYLYPEOLKMTVVVKLISHRECQPHYVGVSEVTTKMLCAADPQWKTDSCQDSSGGPL 380
 QY 353 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
 DB 381 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431
 RESULT 12
 AAP92119
 ID AAP92119 standard; protein; 431 AA.
 XX
 AC AAP92119;
 DT 25-MAR-2003 (updated)
 DT 29-JUN-1990 (first entry)
 XX
 DE Natural human prourokinase.
 XX Human prourokinase; antithrombotic; derivative.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT Misc-difference 1 /note="Optional in new deriv."
 FT Misc-difference 2..155 /note="Incorporated into new deriv."
 FT Misc-difference 135 /note="May be replaced by a non-basic AA in new deriv."
 FT Misc-difference 156 /note="Undefined residue in new deriv."
 FT Misc-difference 157 /note="Pro, Gly, Ala or Val in new deriv."
 FT Misc-difference 158 /note="Lys or Arg in new deriv."
 FT Misc-difference 159 /note="Lys or Arg in new deriv."
 XX WO9901513-A.
 XX 23-FEB-1989.
 XX 18-AUG-1988; 88WO-JP00815.
 XX 19-AUG-1987; 87JP-0204149.
 XX (SAGA) SAGAMI CHEM RES CENTRE.
 XX (CENG) CENTRAL GLASS CO LTD.
 XX (HODO) HODOGAYA CHEM KK.
 XX (NIPS) NIPPON SODA CO.
 XX (NISC) NISSAN CHEM IND LTD.

```
XX Kobayashi Y, Omori M, Yamada C;
XX WPI; 1989-068869/09.
DR N-PSDB; AAN91075.
XX
XX Antithrombotic fast-acting pro-urokinase deriv. -
PT produced by culture of E. coli transformant contg. new plasmid
PT of PMUT9Q family.
XX
XX Disclosure; Fig 1; 75pp; Japanese.
XX
XX A human prourokinase (PU) deriv. is new which is based upon residues
CC 2-155 of natural human prourokinase. The new deriv. is produced by
CC E. coli J103/PMUT9Q-RPK in culture. It is a fast-acting drug for
CC the treatment and prevention of thrombosis.
CC (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 10; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASDTMGPRCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 81 KASDTMGPRCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQGCKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 172
DB 141 PLVQECMVHDCADGKPPPEELKFCQGCKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 200
QY 173 RGSVTVYVCGSLSPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLI 232
DB 201 RGSVTVYVCGSLSPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLI 260
QY 233 LHKDYSADTLAHNDIALKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGK 292
DB 261 LHKDYSADTLAHNDIALKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGK 320
QY 321 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDSSGGL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431
RESULT 13
AAR04253
ID AAR04253 standard; protein; 431 AA.
XX
AC AAR04253;
XX
DT 25-MAR-2003 (updated)
DT 12-SEP-1990 (first entry)
XX
XX Human pro-urokinase from the cDNA of clone pcUK176.
XX
XX Non-glycosylated; pro-urokinase; E. coli; P1rp promoter; MS-2 RBS.
XX
XX Synthetic.
XX
XX EP365894-A.
XX
XX 02-MAY-1990.
XX
```

```
PF 06-OCT-1989; 89EP-0118586.
XX
XX 11-OCT-1988; 88GB-0023833.
XX
XX (FARM ) FARMITALIA ERBA SPA CARLO.
XX
XX Brandazza A, Sarmientos P, Orsini G;
XX WPI; 1990-133447/18.
DR N-PSDB; AAQ04107.
XX
XX Non-glycosylated pro-urokinase prodn. - using E.coli B strains and E.coli
PT promoter P1rp and Shine-Dalgarno sequence MS-2.
XX
XX Disclosure; Page ?; ?pp; English.
XX
XX SER residue at position 21 is the start of the mature prouk.
CC Non-glycosylated prouk (MW 45KD) produced by E.coli B strain containing
CC the sequence..
CC See also AAQ04101-07.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 11; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASDTMGPRCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 81 KASDTMGPRCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQGCKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 172
DB 141 PLVQECMVHDCADGKPPPEELKFCQGCKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 200
QY 173 RGSVTVYVCGSLSPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLI 232
DB 201 RGSVTVYVCGSLSPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLI 260
QY 233 LHKDYSADTLAHNDIALKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGK 292
DB 261 LHKDYSADTLAHNDIALKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDSSGGL 352
DB 321 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDSSGGL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431
RESULT 14
AAR07112
ID AAR07112 standard; protein; 431 AA.
XX
AC AAR07112;
XX
DT 25-MAR-2003 (updated)
DT 24-JAN-1991 (first entry)
XX
XX Human pro-urokinase encoded by plasmid pUK1.
XX
XX pro-urokinase; transgenic mice.
XX
XX Synthetic.
XX
XX EP390592-A.
XX
```

XX 03-OCT-1990.
 PD 30-MAR-1990; 90EP-0303445.
 PF 31-MAR-1989; 89JP-0078574.
 XX (KYOW) KYOMA HAKKO KOGYO KK.
 PA (EXPE-) CENT INST EXPR ANIMALS.
 PA (JIKK-) JIKKEN DOBUTSU CHUO KENK.
 XX Sekine S, Ito S, Katsuki M;
 PI WPI; 1990-299492/40.
 XX N-PSDB; AAQ06049.
 DR Prodn. of recombinant protein, esp. human pro-urokinase - from
 PT milk of transgenic animals using promoter of bovine alpha S1 casein
 PT chromosomal gene.
 XX Example; Table 1; 55pp; English.
 PS E.coli strain C600SF8 was transformed with recombinant plasmid
 CC containing ds DNA derived from human pharynx cancer cell strain
 CC Detroit 562. 10000 colonies were screened and one
 CC positive clone was identified. Plasmid pUK1 was isolated and found
 CC to contain the coding region and 3' non-coding region of pro-UK
 CC downstream of Cys(41). Four silent substitutions were identified
 CC c.f. Holmes et al., Biotechnology, vol.3, p.923 (1985) as follows:
 CC (sic) Asn(254), AAC to AAT; Leu(340), CTA to CTG; Pro(345), CCC to
 CC CCA; Gln(346), CAA to CAG.
 CC See also AAQ06045-Q06048 and AAQ06392.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX SQ Sequence 431 AA;

Query Match 99.4%; Score 2243; DB 11; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
 DB 21 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120
 DB 81 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGK-----LKFCQCKTLRPRFKIIGGEFTTIENQPFAAIYRRH 172
 DB 141 PLVQECMVHDCADGKPKSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFAAIYRRH 200
 QY 173 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
 DB 201 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 233 LHKDYSADTLAHDNDIALKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGK 292
 DB 261 LHKDYSADTLAHDNDIALKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGK 320
 QY 293 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYYSSEVTTKMLCAADPQWKTDSCQDSGGPL 352
 DB 321 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYYSSEVTTKMLCAADPQWKTDSCQDSGGPL 380
 QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
 DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 15
 AAR47903
 ID AAR47903 standard; Protein; 431 AA.
 XX

AC AAR47903;
 XX 13-JUL-1994 (first entry)
 DT Pro-urokinase derivative.
 DE Pro-urokinase derivative.
 XX Pro-urokinase; half-life; thrombolytic; thrombosis; fibrinolytic;
 KW factor.
 XX Homo sapiens.
 OS JP05336965-A.
 PN 21-DEC-1993.
 XX 17-OCT-1991; 91JP-0269615.
 PF 17-OCT-1991; 91JP-0269615.
 XX 17-OCT-1991; 91JP-0269615.
 PR (KYOW) KYOMA HAKKO KOGYO KK.
 XX WPI; 1994-030907/04.
 DR N-PSDB; AAQ55772.
 DR Novel human pro-urokinase derivs. having long half-life - with
 XX high thrombolytic activity, useful for treatment of thrombosis
 PT Disclosure; Page 15-17; 29pp; Japanese.
 PS Sequences (AAQ55771-72) are pro-urokinase derivatives. The products
 CC have an inserted sugar moiety having an amino acid substituted,
 CC depleted or inserted variant around the thrombin cleavage site.
 CC They also have a long half-life allowing them to be used in the
 CC treatment of thrombosis.
 XX SQ Sequence 431 AA;

Query Match 99.4%; Score 2243; DB 15; Length 431;
 Best Local Similarity 98.1%; Pred. No. 1.2e-174;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
 DB 21 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120
 DB 81 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGK-----LKFCQCKTLRPRFKIIGGEFTTIENQPFAAIYRRH 172
 DB 141 PLVQECMVHDCADGKPKSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFAAIYRRH 200
 QY 173 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
 DB 201 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 233 LHKDYSADTLAHDNDIALKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGK 292
 DB 261 LHKDYSADTLAHDNDIALKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGK 320
 QY 293 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYYSSEVTTKMLCAADPQWKTDSCQDSGGPL 352
 DB 321 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYYSSEVTTKMLCAADPQWKTDSCQDSGGPL 380
 QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
 DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 16
 AAR63141
 ID AAR63141 standard; Protein; 431 AA.

```
XX AAR63141;
XX
XX 25-MAR-2003 (updated)
DT 09-JUN-1995 (first entry)
XX
XX Full length human urokinase protein.
XX
XX Human urokinase glycoproteins; cardiovascular diseases;
KW pulmonary embolism.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Sig peptide 1..20
XX Disulfide-bond 70..151
XX Disulfide-bond 91..133
XX Disulfide-bond 122..146
XX Disulfide-bond 168..239
XX Disulfide-bond 209..225
XX Disulfide-bond 217..288
XX Disulfide-bond 313..382
XX Disulfide-bond 345..361
XX Disulfide-bond 372..400
XX Cleavage-site 179..180
XX /note= "cleavage of this site produces a bioactive
XX two chain form of urokinase"
XX
XX EP620279-A1.
XX
XX 19-OCT-1994.
XX
XX 14-APR-1983; 94EP-0104777.
XX
XX 15-APR-1982; 82US-0368773.
XX 14-MAR-1983; 83US-0474930.
XX 14-APR-1983; 83EP-0103629.
XX
XX (GETH ) GENENTECH INC.
XX
XX Heyneker HL, Holmes WE, Vehar GA;
XX WPI; 1994-318362/40.
XX N-PSDB; AAQ73483.
XX
XX Prodn. of human urokinase glycoproteins - using a recombinant
XX expression system used for the treatment of vascular diseases or
XX conditions.
XX
XX Claim 1; Fig 4; 41pp; English.
XX
XX AAQ73483 is the cDNA sequence which encodes AAR63141 the full length
XX 54000 dalton human urokinase (UK) protein. This cDNA was used in
XX the construction of a plasmid capable of transforming either yeast
XX or vertebrate cells, enabling them to produce the 54000 dalton
XX human UK protein. The UK glycoprotein produced could then be used
XX in the treatment of cardiovascular diseases, including pulmonary
XX embolism. The UK produced using this method had the advantage of a
XX specific activity towards fibrin and extant thrombi, not
XX demonstrated previously with UK isolated from natural sources.
XX (Updated on 25-MAR-2003 to correct FN field.)
XX (Updated on 25-MAR-2003 to correct PP field.)
XX (Updated on 25-MAR-2003 to correct PR field.)
XX
XX SQ Sequence 431 AA;
XX
XX Query Match 99.4%; Score 2243; DB 15; Length 431;
XX Best Local Similarity 98.1%; Pred. No. 1.2e-174;
XX Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
XX
XX 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFTGGQHCIDKSKTCYEGNGHYRG 60
XX 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFTGGQHCIDKSKTCYEGNGHYRG 80
```

```
QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCWYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCWYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRH 172
DB 141 PLVQECMVHDCADGKPKSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRH 200
QY 173 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDIYVYVGRSRLNSNTQGENMKFEVENLI 232
DB 201 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDIYVYVGRSRLNSNTQGENMKFEVENLI 260
QY 233 LHKDYSADTLAHNDIALIKIRSKGRCQAQPSRTTQITCLPSMYNDPFGTSCETITGFGK 292
DB 261 LHKDYSADTLAHNDIALIKIRSKGRCQAQPSRTTQITCLPSMYNDPFGTSCETITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLIISHRECOOPHYGSEVTTKMLCAADPQWKTDSCQSGSGGPL 352
DB 321 ENSTDYLYPEQLKMTVVKLIISHRECOOPHYGSEVTTKMLCAADPQWKTDSCQSGSGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYVTVRSVSHFLPWIRSHTKBENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYVTVRSVSHFLPWIRSHTKBENGLAL 431
```

RESULT 17

AA99591

ID AA99591 standard; protein; 431 AA.

XX AA99591;

XX AC AA99591;

XX DT 13-SEP-2000 (first entry)

XX DE Human plasminogen activator urokinase, u-PA.

XX KW Human; serine protease; plasminogen activator; cardiant;

XX KW thrombolytic; heart attack; stroke; blood clotting disorder.

XX OS Homo sapiens.

XX PN WO200032759-A1.

XX PD 08-JUN-2000.

XX PF 06-MAY-1999; 99WO-US09991.

XX PR 02-DEC-1998; 98US-0110588.

XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.

XX PI Lin X, Zhang XC, Tang JJJ;

XX DR WPI; 2000-422975/36.

XX PT Polypeptide with plasminogen activator activity useful as thrombolytic

XX PT agent for treating blood clot disorders e.g. heart attack, comprises 10

XX PT amino acid peptide fragment for recognition or activation of

XX PT plasminogen

XX PS Disclosure; Page 26-28; 41pp; English.

XX CC The present sequence is human plasminogen activator urokinase (u-PA), a

XX CC serine protease which hydrolyses a peptide bond in human plasminogen to

XX CC convert it to its active form, plasmin. Plasminogen is the principal

XX CC serine protease zymogen in the extracellular fluids of vertebrates.

XX CC Plasmin is implicated in pericellular proteolysis associated with a

XX CC wide range of physiological and pathological processes. Plasminogen

XX CC activators regulate plasminogen expression either by hydrolysing a

XX CC peptide bond, as in the case of u-PA, or by forming tight binding

XX CC complexes with plasminogen to spontaneously convert it to plasmin. Review

XX CC of sequence homologues of several plasminogen activators and chymotrypsin

XX CC has identified a six amino acid peptide involved in plasminogen

CC activation. This peptide is particularly useful when inserted between
CC amino acid residues 644 and 645 of full length human plasminogen. Novel
CC plasminogen activators have been made based upon the plasminogen
CC activation/recognition site of plasminogen binding proteins. The
CC polypeptides are useful in preparing thrombolytic agents for treating
CC blood clotting disorders such as heart attack.
XX Sequence 431 AA;
SQ Query Match 99.4%; Score 2243; DB 21; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFOCGOKTLRPRFKIIGGEFTTIENOPWFAAIYRRH 172
DB 141 PLVQECMVHDCADGKSPPEELKFOCGOKTLRPRFKIIGGEFTTIENOPWFAAIYRRH 200
QY 173 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLI 232
DB 201 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHNDIALKIRSKGRCAPSRITQICLPSMYNDPQGTSCIEITGFGK 292
DB 261 LHKDYSADTLAHNDIALKIRSKGRCAPSRITQICLPSMYNDPQGTSCIEITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLISHRECQOPHYVGSSEVTTKMLCAADPQWKTDSCQDGGPL 352
DB 321 ENSTDYLYPEQLKMTVVKLISHRECQOPHYVGSSEVTTKMLCAADPQWKTDSCQDGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431
RESULT 18
AAY50869 ID AAY50869 standard; protein; 431 AA.
XX AC AAY50869;
XX DT 24-FEB-2000 (first entry)
XX DE Human urokinase protein fragment.
XX KW Urokinase; human; thrombolytic agent; streptokinase; antigenic;
XX KW blood clot; heart attack; treatment.
XX OS Homo sapiens.
XX PN WO9957251-A2.
XX PD 11-NOV-1999.
XX PF 06-MAY-1999; 99WO-US10086.
XX PR 06-MAY-1998; 98US-0084392.
XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX PI Zhang XC, Lin X, Tang JUN;
XX DR WPI; 2000-052966/04.
XX PT New thrombolytic agents derived from modified humanized streptokinase,
XX useful for treating blood clot disorders -

XX Disclosure; Page 46-48; 55pp; English.
PS This invention describes a novel thrombolytic agent comprising
CC streptokinase where at least one nonessential portion has been modified.
CC The invention also describes a method of forming a thrombolytic agent
CC comprising determining a nonessential portion of streptokinase and
CC modifying the nonessential portion to render the resulting protein less
CC antigenic. The modified streptokinase is used to treat blood clot
CC disorders, such as heart attacks. The modified streptokinase has less
CC antigenicity than streptokinase but is still able to complex plasminogen
CC and lead to plasminogen activation. Modified streptokinase with the
CC nonessential portions removed or truncated simplify the molecule. Such
CC smaller proteins are cheaper and easier to produce. This sequence
CC represents a fragment of the human urokinase protein which is used in
CC the description of the method of the invention.
XX Sequence 431 AA;
SQ Query Match 99.4%; Score 2243; DB 21; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFOCGOKTLRPRFKIIGGEFTTIENOPWFAAIYRRH 172
DB 141 PLVQECMVHDCADGKSPPEELKFOCGOKTLRPRFKIIGGEFTTIENOPWFAAIYRRH 200
QY 173 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLI 232
DB 201 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHNDIALKIRSKGRCAPSRITQICLPSMYNDPQGTSCIEITGFGK 292
DB 261 LHKDYSADTLAHNDIALKIRSKGRCAPSRITQICLPSMYNDPQGTSCIEITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLISHRECQOPHYVGSSEVTTKMLCAADPQWKTDSCQDGGPL 352
DB 321 ENSTDYLYPEQLKMTVVKLISHRECQOPHYVGSSEVTTKMLCAADPQWKTDSCQDGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431
RESULT 19
AAB84605 ID AAB84605 standard; Protein; 431 AA.
XX AC AAB84605;
XX DT 05-SEP-2001 (first entry)
XX DE Amino acid sequence of urokinase plasminogen activator.
XX KW Growth factor; protein inhibitor; protease; damaged tissue;
XX KW platelet-derived growth factor; PDGF; fibroblast growth factor; FGF;
XX KW connective tissue derived growth factor; CTGF; chrysalin; VEGF;
XX KW keratinocyte-derived growth factor; KGF; epidermal growth factor; EGF;
XX KW transforming growth factor-beta; TGF-beta; matrix metalloproteinase; MMP;
XX KW granulocyte macrophage colony stimulating factor; GM-CSF; uPA;
XX KW vascular endothelial growth factor; urokinase plasminogen activator;
XX KW dermal ulcer; wound.
XX OS Homo sapiens.

PN WO200149309-A2.
XX 12-JUL-2001.
PD 21-DEC-2000; 2000WO-IB01935.
XX 29-DEC-1999; 99GB-0030768.
XX (PFIZ) PFIZER LTD.
XX (PFIZ) PFIZER INC.
XX Davies MJ, Huggins JP, McIntosh PS, Occleston NL;
XX WPI; 2001-418351/44.
XX N-PSDB; AAH28220.
XX Composition for the treatment of damaged tissue i.e. chronic wounds and
PT dermal ulcers comprises an inhibitor agent i.e. a protease and a growth
PT factor .
XX Disclosure; Page 550; 572pp; English.
XX The specification describes a pharmaceutical composition, comprising
XX a growth factor, an inhibitor agent, i.e. a protease. The inhibitor
XX agent inhibits the action of at least one specific adverse protein,
XX i.e. a protease, that is upregulated in a damaged tissue such as a
XX wound environment. Growth factors which are included in the composition
XX of the invention are platelet-derived growth factor (PDGF), fibroblast
XX growth factor (FGF), connective tissue derived growth factor (CTGF),
XX keratinocyte-derived growth factor (KGF), transforming growth
XX factor-beta (TGF-beta), granulocyte macrophage colony stimulating factor
XX (GM-CSF), epidermal growth factor (EGF), vascular endothelial growth
XX factor (VEGF), and chrysalin. Inhibitors which are included in the
XX composition of the invention include inhibitors of urokinase-type
XX plasminogen activator (uPA) and matrix metalloproteinase (MMP). The
XX composition is useful for the treatment of chronic damaged tissue, i.e.
XX wounds and dermal ulcers. The present sequence represents a human uPA,
XX and is used to produce the composition of the invention.
XX Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 22; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQVPSNCDLNGTCTVSNKYFNSIHWCNPKXFGGQHCIDSKTCYEGNGHYRG 60
DB 21 SNELHQVPSNCDLNGTCTVSNKYFNSIHWCNPKXFGGQHCIDSKTCYEGNGHYRG 80
QY 61 KASDTMTGRPCLPWNSATVLQQTVAHRS DALQLGLGKHNYCRPNRRPWCYVQVGLK 120
DB 81 KASDTMTGRPCLPWNSATVLQQTVAHRS DALQLGLGKHNYCRPNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFGCGQKTLRPRFKIIGGFEFTTIENQFPAAIYRRH 172
DB 141 PLVQECMVHDCADGKSPPEELKFGCGQKTLRPRFKIIGGFEFTTIENQFPAAIYRRH 200
QY 173 RGGSVTVVCGSLTSPCWVISAHCFTDYDPKKEDYIVLGRSLNSNTQGMKFEVENLI 232
DB 201 RGGSVTVVCGSLTSPCWVISAHCFTDYDPKKEDYIVLGRSLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNPDQFQTSCEITGFGK 292
DB 261 LHKDYSADTLAHHNDIALKIRSKGRCAPSRITQICLPSMYNPDQFQTSCEITGFGK 320
QY 293 ENSTDYLYPBLQMTVVKLLSHRECQPHYVGGSEVTTKMLCAADPQWKTSCOGDSGGPL 352
DB 321 ENSTDYLYPBLQMTVVKLLSHRECQPHYVGGSEVTTKMLCAADPQWKTSCOGDSGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWRSHRSHKEENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWRSHRSHKEENGLAL 431

RESULT 20
RAG79460
ID AAG79460 standard; Protein; 431 AA.
XX AC AAG79460;
XX 15-NOV-2002 (first entry)
XX DE SC-UPA.
XX Single chain prepro-urokinase; sc-uPA; pro-urokinase; HIV;
XX high molecular weight urokinase-type plasminogen activator; HMW-uPA;
XX long A; long B; EGF-like domain; kringle domain; urokinase receptor;
XX low molecular weight urokinase-type plasminogen activator; LMW-uPA;
XX CD87; binding domain.
XX Homo sapiens.
XX Location/Qualifiers
FH Key 1..20
FT Peptide /note= "Signal peptide"
FT Protein 21..431
FT /label= pro-urokinase/HMW-uPA
FT /note= "With a cleavage between amino acids 178-179"
FT Protein 156..431
FT /label= LMW-uPA
FT /note= "With a cleavage between amino acids 178-179"
XX EP1232755-A2.
XX 21-AUG-2002.
XX 15-FEB-2002; 2002EP-0003555.
XX 20-FEB-2001; 2001JP-0042655.
XX 19-JUN-2001; 2001JP-0184284.
XX (JCRP-) JCR PHARM CO LTD.
XX Wada M, Wada N;
XX WPI; 2002-610512/66.
XX N-PSDB; ABA00207.
XX Anti-HIV agents, comprises ligand molecule that binds to CD87, e.g.
PT high molecular weight urokinase-type plasminogen activator,
PT amino-terminal fragment or an anti-CD87 antibody
XX Disclosure; Page 20-23; 38pp; English.
XX This sequence represents single chain prepro-urokinase (sc-uPA).
XX Pro-urokinase (amino acids 21-431) with a cleavage between amino
XX acids 178 and 179 gives high molecular weight urokinase-type
XX plasminogen activator (HMW-uPA). HMW-uPA is a protein consisting
XX of two peptide chains linked by a di-sulphide bond. The chains,
XX long A and B, are formed by enzymatic cleavage between amino acids
XX 178 and 179 of pro-urokinase. HMW-uPA includes an EGF-like domain,
XX a kringle domain and a urokinase receptor (CD87) binding domain.
XX HMW-uPA is then cleaved between amino acids 155 and 156 to give low
XX molecular weight urokinase-type plasminogen activator (LMW-uPA)
XX (amino acids 156-178 and 179-431), that has no plasminogen activator
XX activity. sc-uPA, or fragments of it, may be used in the anti-HIV
XX agents of the invention which comprise a ligand molecule that binds to
XX CD87. The agents are useful for treating HIV-infected humans for
XX suppression of reproduction of HIV. The anti-HIV agents act by a
XX mechanism of action different from those of conventional drugs, of
XX widening the choice of therapeutics agents and avoiding problems of
XX resistant HIV.
XX Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 23; Length 431;

Best Local Similarity 98.1%; Pred. No. 1.2e-174; Mismatches 0; Indels 8; Gaps 1;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 81 KASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFQCGQKTLRPRFKIIGGFTTIENQPFAPAAIYRRH 172
DB 141 PLVQECMVHDCADGKPPSPPEELKFCQGGKTLRPRFKIIGGFTTIENQPFAPAAIYRRH 200
QY 173 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGEKMFVENLI 232
DB 201 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGEKMFVENLI 260
QY 233 LHKDYSADTLAHNDIALKIRSKGRCAPSPRTIQTICLPSMYNDPQFGTSCIEITGFGK 292
DB 261 LHKDYSADTLAHNDIALKIRSKGRCAPSPRTIQTICLPSMYNDPQFGTSCIEITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGEVTTKMLCAADPQWKTDSCQDGGGGL 352
DB 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGEVTTKMLCAADPQWKTDSCQDGGGGL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENG LAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENG LAL 431

RESULT 21
AAU99228
ID AAU99228 standard; Protein; 431 AA.
XX AAU99228;
AC AAU99228;
XX AAU99228;
DT 24-SEP-2002 (first entry)
XX Human plasminogen activator, urokinase (PLAU).
XX Human; Plasminogen activator; urokinase; PLAU; cancer; enzyme;
KW cytostatic; serine protease; thrombolytic disorder; isogene;
KW pulmonary embolism; chromosome 10q24-qter; haplotype; genotype;
XX SNP; single nucleotide polymorphism; thrombolytic; gene therapy.
XX Homo sapiens.
OS
XX WO200240503-A2.
PN
XX 23-MAY-2002.
PD
XX 14-NOV-2001; 2001WO-US44001.
XX
XX 17-NOV-2000; 2000US-249703P.
XX
XX (GENA-) GENAISSANCE PHARM INC.
PA
XX Anastasio AE, Bentivegna SC, Koshy B;
PI
XX WPI; 2002-519370/55.
DR
XX N-PSDB; ABK86597, ABK86598.
XX
XX Genetic variants of Plasminogen activator, Urokinase (PLAU) isogenes,
PT useful for improving efficiency and reliability in drug development for
PT treating thrombolytic disorders and cancer -
XX
XX Claim 27; Fig 3; 92pp; English.
PS
XX The invention relates to a polynucleotide comprising a first nucleotide
CC sequence (NS1) comprising a PLAU (plasminogen activator, urokinase,
CC a serine protease) isogene selected from isogenes 1-9 and 11-20 given

CC in the specification, where each isogene comprises the regions of the
CC PLAU gene or cDNA and is further defined by the corresponding sequence of
CC polymorphisms (defining single nucleotide polymorphisms, SNP). Also
CC included are methods of haplotyping/genotyping (and predicting the
CC haplotype/genotype of the PLAU gene of an individual, identifying an
CC association between a trait and at least one haplotype or haplotype pair
CC of the PLAU gene, an isolated oligonucleotide for detecting a
CC polymorphism in the PLAU gene, a recombinant non-human organism
CC polynucleotides of at least 10 base pairs encompassing a polymorphic
CC site, an isolated polymorphic variant PLAU protein or fragment, an
CC isolated monoclonal antibody specific for PLAU, a computer system for
CC storing and analysing polymorphism data for the PLAU gene and a genome
CC anthology for the PLAU gene. PLAU is useful in screening for drugs
CC targeting PLAU that are useful for treating thrombolytic disorders and
CC cancers. The methods are useful for improving the efficiency and
CC reliability of the discovery and development of drugs for treating
CC diseases associated with PLAU activity, in validating PLAU as a drug
CC target and in the design of clinical trials for treating a specific
CC condition of disease associated with PLAU activity. The antibody is
CC useful in diagnostic, prognostic and therapeutic methods. PLAU
CC polynucleotides are useful in studying the expression and function of
CC PLAU, and in expressing PLAU protein for use in screening for candidate
CC drugs to treat diseases related to PLAU activity. The gene for PLAU
CC is located on chromosome 10q24-qter. The present sequence represents the
CC PLAU protein.
XX
SQ

Sequence 431 AA;

Query Match 99.4%; Score 2243; DB 23; Length 431;

Best Local Similarity 98.1%; Pred No. 1.2e-174;

Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 81 KASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFQCGQKTLRPRFKIIGGFTTIENQPFAPAAIYRRH 172
DB 141 PLVQECMVHDCADGKPPSPPEELKFCQGGKTLRPRFKIIGGFTTIENQPFAPAAIYRRH 200
QY 173 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGEKMFVENLI 232
DB 201 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGEKMFVENLI 260
QY 233 LHKDYSADTLAHNDIALKIRSKGRCAPSPRTIQTICLPSMYNDPQFGTSCIEITGFGK 292
DB 261 LHKDYSADTLAHNDIALKIRSKGRCAPSPRTIQTICLPSMYNDPQFGTSCIEITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGEVTTKMLCAADPQWKTDSCQDGGGGL 352
DB 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGEVTTKMLCAADPQWKTDSCQDGGGGL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENG LAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENG LAL 431

RESULT 22

AAE17128

ID AAE17128 standard; Protein; 431 AA.

XX AAE17128;

AC AAE17128;

XX 18-APR-2002 (first entry)

DT Human uPA protein.

XX Human; cancer; urokinase-type plasminogen activator; uPA; inflammation;

XX

XX

KW Ets-1 transcription factor; N-acetylglucosaminyltransferase V; Gnt-V;
 KW matrix-type metalloproteinase; MMP-1; MMP-3; gene therapy.
 XX Homo sapiens.

OS

PN WO200196506-A2.

XX

PD 20-DEC-2001.

XX

PF 14-JUN-2001; 2001WO-US19248.

XX

PR 14-JUN-2000; 2000US-0593488.

XX

PA (NYXI-) NYXIS NEURO THERAPIES INC.

XX

PI Yamamoto H, Kroes R, Moskal JR;

XX

XX WPI; 2002-1130746/17.

DR N-PSDB; AAD27855.

XX

XX Identifying a compound for treating cancer, comprises detecting

PT transcription factor Ets-1, N-acetylglucosaminyltransferase V,

PT urokinase-type plasminogen activator, matrix-type metalloproteinase-1

PT and -3 gene expression

XX

XX Example 1; Page 62-63; 63pp; English.

XX

XX The invention relates to a method of identifying a compound for treating

CC cancer. The method involves detecting the expression of a panel of

CC sequences selected from transcription factor Ets-1, urokinase-type

CC plasminogen activator (uPA), N-acetylglucosaminyltransferase V (Gnt-V),

CC matrix-type metalloproteinase (MMP)-1 and MMP-3 in the cell. The method

CC is useful for identifying a compound that affects a cell, particularly a

CC cancer cell or glioma cell, or a cell that is involved in inflammation.

CC It is used for diagnosing and/or treating cancer or other conditions that

CC are affected by one or more members of a panel of genes or their protein

CC product. The method is also useful for drug discovery, drug safety

CC evaluations and in gene therapy. The present sequence is human uPA

CC protein.

XX

XX SQ Sequence 431 AA;

Query Match 99.4%; Score 2243; DB 23; Length 431;

Best Local Similarity 98.1%; Pred. No. 1.2e-174;

Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 SNELHQPNSCDLNGTCTVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 60

Db 21 SNELHQPNSCDLNGTCTVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 80

Qy 61 KASDTMTGRPCLPNSATVLTQTYHAHRSALQLGLGKHNYCRPNDRRPWCYVQVGLK 120

Db 81 KASDTMTGRPCLPNSATVLTQTYHAHRSALQLGLGKHNYCRPNDRRPWCYVQVGLK 140

Qy 121 PLVQECMVHDCADGK-----LKFGCGKTLRPFKIIGGFTTIENQFWFAAAYRRH 172

Db 141 PLVQECMVHDCADGKSPPEELKFCGKTLRPFKIIGGFTTIENQFWFAAAYRRH 200

Qy 173 RGGSVTVVCGSLTSPCWVTSATHCFIDYPKEDYIVVLGRSLNSNTQEMKFEVENLI 232

Db 201 RGGSVTVVCGSLTSPCWVTSATHCFIDYPKEDYIVVLGRSLNSNTQEMKFEVENLI 260

Qy 233 LHKDYSADTLAHHNDIALLKIRSEKRCAPSRITQICLPFSMYNDPQFTSCEITGFGK 292

Db 261 LHKDYSADTLAHHNDIALLKIRSEKRCAPSRITQICLPFSMYNDPQFTSCEITGFGK 320

Qy 293 ENSTDYLYPEOLKMTVVKLISHRECOQPHYVVGSEVTTKMLCAADPOWKTDSCOGDSGGL 352

Db 321 ENSTDYLYPEOLKMTVVKLISHRECOQPHYVVGSEVTTKMLCAADPOWKTDSCOGDSGGL 380

Qy 353 VCSLQGRMTLTGIVSWGRGKALKDPGVYTRVSHFLPWRSHTKKEENGLAL 403

Db 381 VCSLQGRMTLTGIVSWGRGKALKDPGVYTRVSHFLPWRSHTKKEENGLAL 431

RESULT 23

ABU56547

ID ABU56547 standard; Protein; 431 AA.

XX

AC ABU56547;

XX

DT 02-APR-2003 (first entry)

XX

DE Lung cancer-associated polypeptide #140.

XX

KW Lung cancer-associated polypeptide; cytostatic; emphysema;

KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;

KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;

KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;

KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX

OS Unidentified.

XX

PN WO200286443-A2.

XX

PD 31-OCT-2002.

XX

PF 18-APR-2002; 2002WO-US12476.

XX

PR 18-APR-2001; 2001US-284770P.

PR

PR 10-MAY-2001; 2001US-290492P.

PR

PR 09-NOV-2001; 2001US-339245P.

PR

PR 13-NOV-2001; 2001US-350666P.

PR

PR 29-NOV-2001; 2001US-334370P.

PR

PR 12-APR-2002; 2002US-372246P.

XX

XX (EOSB-) EOS BIOTECHNOLOGY INC.

PA

PI Aziz N, Murray R;

XX

XX WPI; 2003-093161/08.

DR N-PSDB; ABX76275.

XX

PT Detecting a lung cancer-associated transcript in a cell from a patient

PT for treating lung cancer, by contacting a biological sample from the

PT patient with a polynucleotide that exhibits increased or decreased

XX expression in lung cancer

XX Claim 27; Page 296; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated

CC transcript in a cell from a patient, comprising contacting a biological

CC sample from the patient with a polynucleotide that selectively hybridises

CC to a sequence that is at least 80 % identical to a gene that exhibits

CC increased or decreased expression in lung cancer samples. Lung

CC cancer-associated polynucleotides and polypeptides are used for

CC identifying a compound that modulates a lung cancer-associated

CC polypeptide, for inhibiting proliferation of a lung cancer-associated

CC cell to treat lung cancer in a patient and for treating a mammal having

CC lung cancer by administering a modulatory compound identified. The

CC methods are useful for treating lung cancer, such as small cell lung

CC cancer, non-small cell lung cancer or other benign or precancerous

CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive

CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial

CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides

CC and polypeptides are useful for diagnostic purposes and as targets for

CC screening for therapeutic compounds that modulate lung cancer, such as

CC antibodies. Sequences ABU56408-ABU56745 represent lung cancer-associated

CC polypeptides of the invention.

XX

XX SQ Sequence 431 AA;

Query Match 99.4%; Score 2243; DB 24; Length 431;

Best Local Similarity 98.1%; Pred. No. 1.2e-174;

Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 SNELHQPNSCDLNGTCTVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 60

Db 21 SNELHQPNSCDLNGTCTVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 80

Qy 61 KASDTMTGRPCLPNSATVLTQTYHAHRSALQLGLGKHNYCRPNDRRPWCYVQVGLK 120

Db 81 KASDTMTGRPCLPNSATVLTQTYHAHRSALQLGLGKHNYCRPNDRRPWCYVQVGLK 140

Qy 121 PLVQECMVHDCADGK-----LKFGCGKTLRPFKIIGGFTTIENQFWFAAAYRRH 172

Db 141 PLVQECMVHDCADGKSPPEELKFCGKTLRPFKIIGGFTTIENQFWFAAAYRRH 200

Qy 173 RGGSVTVVCGSLTSPCWVTSATHCFIDYPKEDYIVVLGRSLNSNTQEMKFEVENLI 232

Db 201 RGGSVTVVCGSLTSPCWVTSATHCFIDYPKEDYIVVLGRSLNSNTQEMKFEVENLI 260

Qy 233 LHKDYSADTLAHHNDIALLKIRSEKRCAPSRITQICLPFSMYNDPQFTSCEITGFGK 292

Db 261 LHKDYSADTLAHHNDIALLKIRSEKRCAPSRITQICLPFSMYNDPQFTSCEITGFGK 320

Qy 293 ENSTDYLYPEOLKMTVVKLISHRECOQPHYVVGSEVTTKMLCAADPOWKTDSCOGDSGGL 352

Db 321 ENSTDYLYPEOLKMTVVKLISHRECOQPHYVVGSEVTTKMLCAADPOWKTDSCOGDSGGL 380

Qy 353 VCSLQGRMTLTGIVSWGRGKALKDPGVYTRVSHFLPWRSHTKKEENGLAL 403

Db 381 VCSLQGRMTLTGIVSWGRGKALKDPGVYTRVSHFLPWRSHTKKEENGLAL 431

Qy 1 SNELHQPVSNCDCNLNGTCSVKNYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 60
Db 21 SNELHQPVSNCDCNLNGTCSVKNYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 80
Qy 61 KASDTMTGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNRRPWCYVQVGLK 120
Db 81 KASDTMTGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNRRPWCYVQVGLK 140
Qy 121 PLVQECMWHDCADGK-----LKFCGQCKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 172
Db 141 PLVQECMWHDCADGKSPPEELKFCGQCKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 200
Qy 173 RGSVTVYVCGSLSPCWVTSATHCFIDYPKEDYIVYLGSRSLNSNTQGEKMFVENLI 232
Db 201 RGSVTVYVCGSLSPCWVTSATHCFIDYPKEDYIVYLGSRSLNSNTQGEKMFVENLI 260
Qy 233 LHKDYSADTLAHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 292
Db 261 LHKDYSADTLAHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 320
Qy 293 ENSTDYLYPEQLKMTVVVKLISHRECQCPHYVYGVSEVTTKMLCAADPQWKTDSCQDSSGGL 352
Db 321 ENSTDYLYPEQLKMTVVVKLISHRECQCPHYVYGVSEVTTKMLCAADPQWKTDSCQDSSGGL 380
Qy 353 VCSLQGRMTLFGIVSWGRCALKDKPGVYTVRSHFLPWIRSHTKENGLAL 403
Db 381 VCSLQGRMTLFGIVSWGRCALKDKPGVYTVRSHFLPWIRSHTKENGLAL 431

RESULT 24
ABU56708
ID ABU56708 standard; Protein; 431 AA.
XX AC ABU56708;
XX DT 02-APR-2003 (first entry)
XX DE Lung cancer-associated polypeptide #301.
XX KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; asthma; bronchiectasis.
XX OS Unidentified.
XX PN WO200286443-A2.
XX PD 31-OCT-2002.
XX PF 18-APR-2002; 2002WO-US12476.
XX PR 18-APR-2001; 2001US-284770P.
XX PR 10-MAY-2001; 2001US-290492P.
XX PR 09-NOV-2001; 2001US-339245P.
XX PR 13-NOV-2001; 2001US-350666P.
XX PR 29-NOV-2001; 2001US-334370P.
XX PR 12-APR-2002; 2002US-372246P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Aziz N, Murray R;
XX DR WPI; 2003-093161/08.
XX DR N-PSDB; ABX76437.
XX PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer
XX OS Claim 27; Page 424; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung
CC cancer-associated polynucleotides and polypeptides are used for
CC identifying a compound that modulates a lung cancer-associated
CC polypeptide, for inhibiting proliferation of a lung cancer-associated
CC cell to treat lung cancer in a patient and for treating a mammal having
CC lung cancer by administering a modulatory compound identified. The
CC methods are useful for treating lung cancer, such as small cell lung
CC cancer, non-small cell lung cancer or other benign or precancerous
CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
CC and polypeptides are useful for diagnostic purposes and as targets for
CC screening for therapeutic compounds that modulate lung cancer, such as
CC antibodies. Sequences ABU56408-ABU56745 represent lung cancer-associated
CC polypeptides of the invention.
XX SQ Sequence 431 AA;
Query Match 99.4%; Score 2243; DB 24; Length 431;
Best Local Similarity 98.1%; Pred. No. 1.2e-174; Indels 8; Gaps 1;
Matches 403; Conservative 0; Mismatches 0;
Qy 1 SNELHQPVSNCDCNLNGTCSVKNYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 60
Db 21 SNELHQPVSNCDCNLNGTCSVKNYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 80
Qy 61 KASDTMTGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNRRPWCYVQVGLK 120
Db 81 KASDTMTGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNRRPWCYVQVGLK 140
Qy 121 PLVQECMWHDCADGK-----LKFCGQCKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 172
Db 141 PLVQECMWHDCADGKSPPEELKFCGQCKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 200
Qy 173 RGSVTVYVCGSLSPCWVTSATHCFIDYPKEDYIVYLGSRSLNSNTQGEKMFVENLI 232
Db 201 RGSVTVYVCGSLSPCWVTSATHCFIDYPKEDYIVYLGSRSLNSNTQGEKMFVENLI 260
Qy 233 LHKDYSADTLAHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 292
Db 261 LHKDYSADTLAHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 320
Qy 293 ENSTDYLYPEQLKMTVVVKLISHRECQCPHYVYGVSEVTTKMLCAADPQWKTDSCQDSSGGL 352
Db 321 ENSTDYLYPEQLKMTVVVKLISHRECQCPHYVYGVSEVTTKMLCAADPQWKTDSCQDSSGGL 380
Qy 353 VCSLQGRMTLFGIVSWGRCALKDKPGVYTVRSHFLPWIRSHTKENGLAL 403
Db 381 VCSLQGRMTLFGIVSWGRCALKDKPGVYTVRSHFLPWIRSHTKENGLAL 431

RESULT 25
ABU11076
ID ABU11076 standard; Protein; 431 AA.
XX AC ABU11076;
XX DT 05-FEB-2003 (first entry)
XX DE Human urokinase plasminogen activator.
XX KW Urokinase plasminogen activator; gene therapy; cancer;
KW hyperproliferative disorder; cancer; breast cancer; colon cancer;
KW bone cancer; brain cancer; ovary cancer; cervix cancer;
KW endometrium cancer; stomach cancer; kidney cancer; tumour metastasis.
XX OS Homo sapiens.


```
QY 353 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
Db 381 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 27
AAR20538
ID AAR20538 standard; Protein; 434 AA.
XX
AC AAR20538;
XX
DT 25-MAR-2003 (updated)
DT 21-MAY-1992 (first entry)
XX
DE Amidated deriv. of pro-urokinase (3).
XX
KW Pro-urokinase; plasminogen activator; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..431
FT /label= pro-UK
XX
XX DE4122688-A.
XX
PD 16-JAN-1992.
XX
PF 09-JUL-1991; 91DE-4122688.
XX
PR 12-JUL-1990; 90GB-0015369.
PR 10-JUL-1991; 91GB-0014846.
XX
XX (FARM ) FARMITALIA ERBA SRL CARLO.
XX
PI Gozzini L, Visco C, Perego R, Roncucci R, Sarmientos P;
DR WPI; 1992-025815/04.
XX
XX New amidated derivs. of human pro-urokinase - are fibrinolytic
PT and can be used to treat acute myocardial infarction, pulmonary
PT embolism or deep venous thrombosis
XX
PS Claim 4,9; Page 8; 18pp; German.
XX
XX The protein has fibrinolytic activity and can be used in the same
CC way as PUK, e.g. for treating acute myocardial infarction, lung
CC embolism and deep venous thrombosis. It has greater affinity for
CC plasminogen bound to fibrin than for circulating plasminogen, so
CC have high selectivity for thrombi with reduced chance of bleeding.
CC Compared with the COOH-terminated cpds., it has a better stability
CC against most carboxypeptidases and prolonged half life.
CC See also AAQ20360, AAQ20754-58 and AAR20536-38.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 434 AA;
XX
Query Match 99.4%; Score 2243; DB 13; Length 434;
Best Local Similarity 98.1%; Pred. No. 1.2e-174;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNIHWCNCPKFGGQHCHBIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFNIHWCNCPKFGGQHCHBIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRCLPNSATVLQOQYHAHRSDALQGLGKHNYCENPNRRPRPCYVOVGLK 120
Db 81 KASTDTMGRCLPNSATVLQOQYHAHRSDALQGLGKHNYCENPNRRPRPCYVOVGLK 140
QY 121 PLVQECWVHDCADGK-----LKFQCGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 172
Db 141 PLVQECWVHDCADGKPKSSPEELKFQCGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 200
```

```
QY 173 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSRLSNNTQGMKFEVENLI 232
Db 201 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSRLSNNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALALKIRSKGRCAPQSPRTIQTICLPSMYNDPQFGTSCBITGFGK 292
Db 261 LHKDYSADTLAHHNDIALALKIRSKGRCAPQSPRTIQTICLPSMYNDPQFGTSCBITGFGK 320
QY 293 ENSTDYLYPEOLKMTVVKLISHRECQOPHYVGGSEVTTKMLCAADPQWKTSCQDSSGGL 352
Db 321 ENSTDYLYPEOLKMTVVKLISHRECQOPHYVGGSEVTTKMLCAADPQWKTSCQDSSGGL 380
QY 353 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
Db 381 VCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 28
AAR20536
ID AAR20536 standard; Protein; 436 AA.
XX
AC AAR20536;
XX
DT 25-MAR-2003 (updated)
DT 21-MAY-1992 (first entry)
XX
DE Amidated deriv. of pro-urokinase (1).
XX
KW Pro-urokinase; plasminogen activator; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..431
FT /label= pro-UK
FT /note= "the amidated form of pro-UK is
FT Misc-difference 433..436 Claimed in claim 1"
FT /note= "may be any amino acid, pref. Lys or Arg,
FT or 0-4 amino acids may be omitted"
XX
XX DE4122688-A.
XX
PD 16-JAN-1992.
XX
PF 09-JUL-1991; 91DE-4122688.
XX
PR 12-JUL-1990; 90GB-0015369.
PR 10-JUL-1991; 91GB-0014846.
XX
XX (FARM ) FARMITALIA ERBA SRL CARLO.
XX
PI Gozzini L, Visco C, Perego R, Roncucci R, Sarmientos P;
DR WPI; 1992-025815/04.
XX
XX New amidated derivs. of human pro-urokinase - are fibrinolytic
PT and can be used to treat acute myocardial infarction, pulmonary
PT embolism or deep venous thrombosis
XX
PS Claim 1,4,7; Page 8; 18pp; German.
XX
XX The protein has fibrinolytic activity and can be used in the same
CC way as PUK, e.g. for treating acute myocardial infarction, lung
CC embolism and deep venous thrombosis. It has greater affinity for
CC plasminogen bound to fibrin than for circulating plasminogen, so
CC have high selectivity for thrombi with reduced chance of bleeding.
CC Compared with the COOH-terminated cpds., it has a better stability
CC against most carboxypeptidases and prolonged half life.
CC See also AAQ20360, AAQ20754-58 and AAR20536-38.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 436 AA;
```

Query Match 99.4%; Score 2243; DB 13; Length 436;
Best Local Similarity 98.1%; Pred. No. 1.2e-174; Mismatches 0; Indels 8; Gaps 1;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNPKFKGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 81 KASDTMTGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQCGQKTLRPFKIIIGGEFTTIENQWPAAYRRH 172
DB 141 PLVQECMVHDCADGKPKSPPEELKFCQCGQKTLRPFKIIIGGEFTTIENQWPAAYRRH 200
QY 173 RGGSVTVYVCGGSLISPCWISATHCFIDYPKKEDIYVYLGSRNSNTQGMKFEVENLI 232
DB 201 RGGSVTVYVCGGSLISPCWISATHCFIDYPKKEDIYVYLGSRNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCETIGFGK 292
DB 261 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCETIGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLIISHRECCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGPL 352
DB 321 ENSTDYLYPEQLKMTVVKLIISHRECCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 29
AAW24578
ID AAW24578 standard; Protein; 430 AA.
XX AC AAW24578;
XX DT 25-MAR-2003 (updated)
XX DT 11-NOV-1997 (first entry)
XX DE Inhibitor resistant urokinase.
XX KW Urokinase; inhibitor resistant; plasminogen; human; whey acid protein;
KW plasminogen activator inhibitor-1; plasminogen activator; protease; WAP;
KW plasmin; proteolytic enzyme; clot lysis; heart attack; heart muscle.
XX OS Homo sapiens.
XX OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..19
FT /note= "whey acid protein signal peptide"
FT Protein 20..430
FT /note= "urokinase"
FT Domain 20..64
FT /note= "E-domain"
FT Domain 61..150
FT /note= "Kringle-1 domain"
FT Domain 179..430
FT /note= "P-domain"
FT Misc-difference 198..203
FT /note= "deleted in modified urokinase of the invention"
XX US5648253-A.
XX 15-JUL-1997.
XX 08-SEP-1992; 92US-0942157.
XX 20-DEC-1990; 90US-0631673.

PR 08-SEP-1992; 92US-0942157.
XX (TSIT-) TSI CORP.
XX Wei C;
XX WPI; 1997-372062/34.
XX N-PSDB; AAT80075.
PT Deletion-modified urokinase protein - with increased resistance to
PT inhibition by plasminogen activator inhibitor-1
XX Disclosure; Column 15-18; 16pp; English.
XX This sequence represents the full length urokinase, including the
CC whey acid protein (WAP) signal peptide. This sequence has residues
CC 179-184 of the urokinase sequence deleted to create the modified
CC urokinase of the invention. The modified urokinase (see AAW24579)
CC cleaves plasminogen, and has a lower binding affinity for plasminogen
CC activator inhibitor-1 than the corresponding unmodified urokinase.
CC Urokinase is one of two types of mammalian plasminogen activators (PA),
CC the other being tissue type PA. PAs catalyze the conversion of the
CC circulating zymogen plasminogen to the broad spectrum protease plasmin by
CC limited proteolysis. The modified urokinase can be used for clot lysis,
CC specifically to dissolve heart attack-causing clots before they cause
CC permanent damage to heart muscle. The urokinase mutant is more resistant
CC to inhibition by plasminogen activator inhibitors than the unmodified
CC urokinase. It can be selectively expressed and secreted from the mammary
CC glands of transgenic animals.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX Sequence 430 AA;
QY Query Match 99.2%; Score 2240; DB 18; Length 430;
DB Best Local Similarity 97.8%; Pred. No. 2.1e-174; Mismatches 1; Indels 8; Gaps 1;
Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
DB 20 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNPKFKGGQHCIEDKSKTCYEGNGHFYRG 79
QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 80 KASDTMTGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 139
QY 121 PLVQECMVHDCADGK-----LKFCQCGQKTLRPFKIIIGGEFTTIENQWPAAYRRH 172
DB 140 PLVQECMVHDCADGKPKSPPEELKFCQCGQKTLRPFKIIIGGEFTTIENQWPAAYRRH 199
QY 173 RGGSVTVYVCGGSLISPCWISATHCFIDYPKKEDIYVYLGSRNSNTQGMKFEVENLI 232
DB 200 RGGSVTVYVCGGSLISPCWISATHCFIDYPKKEDIYVYLGSRNSNTQGMKFEVENLI 259
QY 233 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCETIGFGK 292
DB 260 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCETIGFGK 319
QY 293 ENSTDYLYPEQLKMTVVKLIISHRECCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGPL 352
DB 320 ENSTDYLYPEQLKMTVVKLIISHRECCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGPL 379
QY 353 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 380 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 430

RESULT 30
AAP91886
ID AAP91886 standard; protein; 431 AA.
XX AAP91886;
XX 25-MAR-2003 (updated)

DT 31-OCT-2002 (updated)
DT 16-APR-1990 (first entry)
XX Sequence of prourokinase.
DE Low mol. wt. plasminogen activator; prourokinase; dissolve blood clot.
XX Homo sapiens.
OS
XX Location/Qualifiers
FH Key 21..431
FT Protein /note= "Mature prourokinase."
FT Region 170..179
FT /note= "Preferred initiation region for the low mol. wt.
FT plasminogen activators."
XX
PN EP316058-A.
XX
XX PD 17-MAY-1989.
XX
XX PF 07-OCT-1988; 88EP-0309417.
XX
XX PR 09-OCT-1987; 87US-0107370.
XX PR 27-SEP-1988; 88US-0248727.
XX
XX PA. (COLB) COLLABORATIVE RES INC.
XX
XX PI Mao JI;
XX
XX WPI: 1989-146601/20.
XX N-PSDB; AAN91740.
XX
XX Modified low mol. wt. plasminogen activator- formed of amino acids
PT comprising the amino acid portion of prourokinase from 150 to 411
XX
XX Fig 1; Page -: 27pp; English.
XX
XX A low mol. wt. polypeptide plasminogen activator (PA) is claimed which is
CC formed from amino acids 150-411 of prourokinase. The preferred initiation
CC region for the low mol. wt. PA is indicated (see FT). A low mol wt. PA
CC can be injected into blood in the body in vivo to dissolve clots
CC without harm.
CC (Updated on 31-OCT-2002 to add missing OS field.)
CC (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
XX SQ Sequence 431 AA;

Query Match 99.2%; Score 2240; DB 10; Length 431;
Best Local Similarity 97.8%; Pred. No. 2.1e-174;
Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LXFCQCKTLRPRFKIIGGEFTTIENQPFAAIYRRH 172
DB 141 PLVQECMVHDCADGKSPPEELXFCQCKTLRPRFKIIGGEFTTIENQPFAAIYRRH 200
QY 173 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 232
DB 201 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTVAHNDIALIKIRSKRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGK 292
DB 261 LHKDYSADTVAHNDIALIKIRSKRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLIISHRECQPHYGSEVTTKMLCAADPQWKTDSCQDGGG 352

DB 321 ENSTDYLYPEQLKMTVVKLIISHRECQPHYGSEVTTKMLCAADPQWKTDSCQDGGG 380
QY 353 VCSLQGRMTLTGIVSWGRGKALKDKGVYTRVSHFLPWIRSHTKKEENG 403
DB 381 VCSLQGRMTLTGIVSWGRGKALKDKGVYTRVSHFLPWIRSHTKKEENG 431
Search completed: December 3, 2003, 14:39:14
Job time : 64.482 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:35:43 ; Search time 20.8448 Seconds
(without alignments)
818.010 Million cell updates/sec

Title: US-09-880-503-6

Perfect score: 2257

Sequence: 1 SNEHQVPSNCDCLNGTCTV.....VSHFLPWIRSHTEENGLAL 403

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2243.5	99.4	430	6	5219569-2 Patent No. 5219569
2	2243	99.4	411	1	US-08-087-163-1 Sequence 1, Appl
3	2243	99.4	411	1	US-08-286-748B-18 Sequence 18, Appl
4	2243	99.4	411	1	US-08-153-799-18 Sequence 18, Appl
5	2243	99.4	431	4	US-09-101-272G-1 Sequence 1, Appl
6	2243	99.4	431	6	5188829-1 Patent No. 5188829
7	2240	99.2	430	1	US-07-942-157A-3 Sequence 3, Appl
8	2233	98.9	411	3	US-09-181-816-1 Sequence 1, Appl
9	2219	98.3	411	2	US-08-560-098A-48 Sequence 48, Appl
10	1965	87.1	432	2	US-08-560-098A-47 Sequence 47, Appl
11	1964	87.0	365	1	US-08-093-741-83 Sequence 83, Appl
12	1964	87.0	365	1	US-08-720-012-83 Sequence 83, Appl
13	1964	87.0	393	3	US-08-560-098A-44 Sequence 44, Appl
14	1964	87.0	393	3	US-08-987-024C-24 Sequence 24, Appl
15	1964	87.0	393	3	US-08-987-024C-25 Sequence 25, Appl
16	1466.5	65.0	306	2	US-08-560-098A-45 Sequence 45, Appl
17	1466.5	65.0	331	2	US-08-560-098A-46 Sequence 46, Appl
18	1382	61.2	253	3	US-08-944-483-73 Sequence 73, Appl
19	1374	60.9	254	2	US-08-560-098A-49 Sequence 49, Appl
20	983	43.6	200	4	US-09-101-272G-73 Sequence 73, Appl
21	867.5	38.4	477	2	US-08-560-098A-51 Sequence 51, Appl
22	855.5	38.3	527	6	5520913-1 Patent No. 5520913
23	859.5	38.0	527	6	US-07-609-510B-16 Sequence 16, Appl
24	858.5	38.0	527	5	PCT-US91-01025A-2 Sequence 2, Appl
25	858.5	38.0	527	5	5185259-8 Patent No. 5185259
26	858.5	38.0	527	6	US-08-811-949-43 Sequence 43, Appl
27	858.5	38.0	562	2	US-08-811-949-43 Sequence 43, Appl

28	858.5	38.0	562	2	US-08-560-098A-50 Sequence 50, Appl
29	858.5	38.0	562	2	US-08-883-795A-38 Sequence 38, Appl
30	858.5	38.0	562	6	5185259-3 Patent No. 5185259
31	858.5	38.0	562	6	5200340-2 Patent No. 5200340
32	858.5	38.0	562	6	5344773-2 Patent No. 5344773
33	851.5	37.7	562	6	5244676-5 Patent No. 5244676
34	825	36.6	157	3	US-08-142-590B-25 Sequence 25, Appl
35	808	35.8	355	2	US-08-811-949-59 Sequence 59, Appl
36	803	35.6	437	2	US-08-811-949-51 Sequence 51, Appl
37	800	35.4	437	2	US-08-811-949-57 Sequence 57, Appl
38	799	35.4	208	4	US-09-101-272G-98 Sequence 98, Appl
39	796	35.3	472	2	US-08-811-949-63 Sequence 63, Appl
40	793	35.1	138	2	US-08-797-689-12 Sequence 12, Appl
41	793	35.1	437	2	US-08-811-949-55 Sequence 55, Appl
42	792	35.0	437	2	US-08-811-949-49 Sequence 49, Appl
43	791	35.0	355	2	US-08-811-949-47 Sequence 47, Appl
44	790	35.0	356	1	US-08-427-640-8 Sequence 8, Appl
45	788	34.9	194	4	US-09-101-272G-80 Sequence 80, Appl

ALIGNMENTS

RESULT 1
5219569-2
; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR, GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO: 2
; LENGTH: 430
5219569-2

Query Match 99.4%; Score 2243.5; DB 6; Length 430;
Best Local Similarity 98.3%; Pred. No. 1.8e-188;
Matches 403; Conservative 0; Mismatches 0; Indels 7; Gaps 1;

QY	1	SNEHQVPSNCDCLNGTCTVSNKYFSNIHWCNCPKFGQHCIEIDKSKTCYEGNGHFYRG	60
DB	21	SNEHQVPSNCDCLNGTCTVSNKYFSNIHWCNCPKFGQHCIEIDKSKTCYEGNGHFYRG	80
QY	61	KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRPWCYVOVGLK	120
DB	81	KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRPWCYVOVGLK	140
QY	121	PLVQECMWHDCADGK-----LKFCGQKTLRPRFKIIGGFTTIENQWFAAIYRRHR	173
DB	141	PLVQECMWHDCADGKSPPEELKFCGQKTLRPRFKIIGGFTTIENQWFAAIYRRHR	200
QY	174	GSVTVVCGSLISPCWVISATHCFIDYPKKEDYVYLGSRSLNSQTQCKMFEVENLL	233
DB	201	GSVTVVCGSLISPCWVISATHCFIDYPKKEDYVYLGSRSLNSQTQCKMFEVENLL	260
QY	234	HKDYSADT LAHNDIALKIRSKGRCAQPSRTIOTICLPSMYNDPQFCTSCITGFGKE	293
DB	261	HKDYSADT LAHNDIALKIRSKGRCAQPSRTIOTICLPSMYNDPQFCTSCITGFGKE	320
QY	294	NSTDYLYPEQLKMTVVKLISHRECOPHYGSEVTTKMLCAADPQWKTDCQDGGGLV	353
DB	321	NSTDYLYPEQLKMTVVKLISHRECOPHYGSEVTTKMLCAADPQWKTDCQDGGGLV	380
QY	354	CSLQGRMTLTGIVSWGRGCKALDKPGVYTRVSHFLPWIRSHRTEENGLAL	403
DB	381	CSLQGRMTLTGIVSWGRGCKALDKPGVYTRVSHFLPWIRSHRTEENGLAL	430

RESULT 2

US-08-087-163-1
 ; Sequence 1, Application US/08087163
 ; Patent No. 5472692
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jian-Ning
 ; APPLICANT: Gurewich, Victor
 ; TITLE OF INVENTION: PRO-UKINASE MUTANTS
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; COMPUTER: IBM PS/2 Model 502 or 55SX
 ; OPERATING SYSTEM: MS-DOS (Version 5.0)
 ; SOFTWARE: WordPerfect (Version 5.1)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/087,163
 ; FILING DATE: 07/02/93
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fasse, J. Peter
 ; REGISTRATION NUMBER: 32,983
 ; REFERENCE/DOCKET NUMBER: 04353/003001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 542-5070
 ; TELEFAX: (617) 542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 411
 ; TYPE: amino acid
 ; STRANDEDNESS: N/A
 ; TOPOLOGY: N/A
 ;

Query Match 99.4%; Score 2243; DB 1; Length 411;
 Best Local Similarity 98.1%; Pred. No. 1.9e-188;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
 Db 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
 Qy 61 KASTDTMGRPCLPWN SATVLQOITYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 120
 Db 61 KASTDTMGRPCLPWN SATVLQOITYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 120
 Qy 121 PLVQECMVHCDADGK-----LKFCQCKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
 Db 121 PLVQECMVHCDADGKPPSPPELKFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 180
 Qy 173 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 232
 Db 181 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 240
 Qy 233 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSCITGFGK 292
 Db 241 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSCITGFGK 300
 Qy 293 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGGL 352
 Db 301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGGL 360
 Qy 353 VCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWIRSHKTEENGLAL 403

Db 361 VCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWIRSHKTEENGLAL 411

RESULT 3

US-08-286-748B-18
 ; Sequence 18, Application US/08286748B
 ; Patent No. 5759542
 ; GENERAL INFORMATION:
 ; APPLICANT: Victor Gurewich
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
 ; TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF
 ; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; COMPUTER: IBM PS/2 Model 502 or 55SX
 ; OPERATING SYSTEM: MS-DOS (Version 5.0)
 ; SOFTWARE: WordPerfect (Version 5.1)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/286,748B
 ; FILING DATE: August 5, 1994
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: J. Peter Fasse
 ; REGISTRATION NUMBER: 32,983
 ; REFERENCE/DOCKET NUMBER: 04547/013001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 542-5070
 ; TELEFAX: (617) 542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 18:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 411
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ;

Query Match 99.4%; Score 2243; DB 1; Length 411;
 Best Local Similarity 98.1%; Pred. No. 1.9e-188;
 Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
 Db 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
 Qy 61 KASTDTMGRPCLPWN SATVLQOITYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 120
 Db 61 KASTDTMGRPCLPWN SATVLQOITYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 120
 Qy 121 PLVQECMVHCDADGK-----LKFCQCKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
 Db 121 PLVQECMVHCDADGKPPSPPELKFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 180
 Qy 173 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 232
 Db 181 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 240
 Qy 233 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSCITGFGK 292
 Db 241 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSCITGFGK 300

QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDSDGGPL 352
Db 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDSDGGPL 360
QY 353 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKBENGLAL 403
Db 361 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKBENGLAL 411

RESULT 4
US-08-153-799-18
; Sequence 18, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153.799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-153-799-18

Query Match 99.4%; Score 2243; DB 1; Length 411;
Best Local Similarity 98.1%; Pred. No. 1.9e-188;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 60
Db 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 60
QY 61 KASDTMTGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
Db 61 KASDTMTGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120

QY 121 PLVQECMVHDCADGK-----LKFQCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 172
Db 121 PLVQECMVHDCADGKPPSPPELAFQCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 180
QY 173 RGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVLGRSRLNSNTGEMKFEVENLI 232
Db 181 RGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVLGRSRLNSNTGEMKFEVENLI 240
QY 233 LHKDYSADTLAHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCETITGFGK 292
Db 241 LHKDYSADTLAHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCETITGFGK 300
QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDSDGGPL 352
Db 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDSDGGPL 360
QY 353 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKBENGLAL 403
Db 361 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTKBENGLAL 411

RESULT 5
US-09-101-272G-1
; Sequence 1, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101.272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (21)..()
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (20)..()
; OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)
US-09-101-272G-1

Query Match 99.4%; Score 2243; DB 4; Length 431;
Best Local Similarity 98.1%; Pred. No. 2e-188;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 60
Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 80
QY 61 KASDTMTGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
Db 81 KASDTMTGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFQCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 172
Db 141 PLVQECMVHDCADGKPPSPPELAFQCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 200
QY 173 RGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVLGRSRLNSNTGEMKFEVENLI 232
Db 201 RGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVLGRSRLNSNTGEMKFEVENLI 260
QY 233 LHKDYSADTLAHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCETITGFGK 292
Db 261 LHKDYSADTLAHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCETITGFGK 320

QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYHSEVTTKMLCAADPQWKTDSCQDGGPL 352
DB 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYHSEVTTKMLCAADPQWKTDSCQDGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 6
5188829-1
; Patent No. 5188829
; APPLICANT: KOBAYASHI, YO-ICHI, OMORI, MUNEKI, YAMADA, CHIRAKO
; TITLE OF INVENTION: RAPIDLY ACTING PROUKINASE
; NUMBER OF SEQUENCES: 23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/340,007
; FILING DATE: 18-AUG-1988
; SEQ ID NO: 1:
; LENGTH: 431
5188829-1

Query Match 99.4%; Score 2243; DB 6; Length 431;
Best Local Similarity 98.1%; Pred. No. 2e-188;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQCGOKTLRPRFKIIGGEFTTINQWPFAAIYRRH 172
DB 141 PLVQECMVHDCADGKPPSPPEELKFQCGQKTLRPRFKIIGGEFTTINQWPFAAIYRRH 200
QY 173 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYIVYLGRSRLNSNTQGMKFEVENLI 232
DB 201 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYIVYLGRSRLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCEITGFGK 292
DB 261 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCEITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYHSEVTTKMLCAADPQWKTDSCQDGGPL 352
DB 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYHSEVTTKMLCAADPQWKTDSCQDGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 7
US-07-942-157A-3
; Sequence 3, Application US/07942157A
; Patent No. 5648253
; GENERAL INFORMATION:
; APPLICANT: Wei, Cha-Mer
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,157A
; FILING DATE: 19920908
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631673
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: TS1108Cont.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)815-6508
; TELEFAX: (404)815-6555
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..19
; OTHER INFORMATION: /label= peptide
; OTHER INFORMATION: /note= "WAP signal"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 198..203
; OTHER INFORMATION: /label= modified
; OTHER INFORMATION: /note= "six amino acids deleted in mutant"
US-07-942-157A-3

Query Match 99.2%; Score 2240; DB 1; Length 430;
Best Local Similarity 97.8%; Pred. No. 3.7e-188;
Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCEDKSKTCYEGNGHFYRG 60
DB 20 SNELHQPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCEDKSKTCYEGNGHFYRG 79
QY 61 KASTDTMGRPCLPWNATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 80 KASTDTMGRPCLPWNATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 139
QY 121 PLVQECMVHDCADGK-----LKFCQCGOKTLRPRFKIIGGEFTTINQWPFAAIYRRH 172
DB 140 PLVQECMVHDCADGKPPSPPEELKFQCGQKTLRPRFKIIGGEFTTINQWPFAAIYRRH 199
QY 173 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYIVYLGRSRLNSNTQGMKFEVENLI 232
DB 200 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYIVYLGRSRLNSNTQGMKFEVENLI 259
QY 233 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCEITGFGK 292
DB 260 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCEITGFGK 319
QY 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYHSEVTTKMLCAADPQWKTDSCQDGGPL 352
DB 320 ENSTDYLYPEQLKMTVVKLISHRECOQPHYHSEVTTKMLCAADPQWKTDSCQDGGPL 379
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
DB 380 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 430

RESULT 8
US-09-181-816-1
; Sequence 1, Application US/09181816
; Patent No. 6277818
; GENERAL INFORMATION:

```
; APPLICANT: MAZAR, Andrew P.
; APPLICANT: JONES, Terence R.
; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
; TITLE OF INVENTION: PLASMINOGEN ACTIVATOR RECEPTOR
; FILE REFERENCE: 329042000300 SIDN 1-7
; CURRENT APPLICATION NUMBER: US/09/181.816
; CURRENT FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-181-816-1

Query Match 98.9%; Score 2233; DB 3; Length 411;
Best Local Similarity 97.8%; Pred. No. 1.4e-187;
Matches 402; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

QY 1 SNELHQPSPNCDCNLGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPSPNCDCNLGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK-----LKFCQGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 172
DB 121 LLVQECMVHDCADGKFPSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
QY 173 RGSVTVYCGGSLSPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 232
DB 181 RGSVTVYCGGSLSPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 240
QY 233 LHKDYSADTLAHNNDIALLKIRSKGRCQAPSRITQICLPSMYNDPQFGTSCEITGFGK 292
DB 241 LHKDYSADTLAHNNDIALLKIRSKGRCQAPSRITQICLPSMYNDPQFGTSCEITGFGK 300
QY 293 ENSTDYLYPEQLKMTVVKLIHSHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGPL 352
DB 301 ENSTDYLYPEQLKMTVVKLIHSHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGPL 360
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
DB 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

RESULT 9
US-08-560-098A-48
; Sequence 48. Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEEDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560.098A

; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-48

Query Match 98.3%; Score 2219; DB 2; Length 411;
Best Local Similarity 97.3%; Pred. No. 2.4e-186;
Matches 400; Conservative 0; Mismatches 3; Indels 8; Gaps 1;

QY 1 SNELHQPSPNCDCNLGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPSPNCDCNLGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK-----LKFCQGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 172
DB 121 PLVQECMVHDCADGKFPSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
QY 173 RGSVTVYCGGSLSPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 232
DB 181 RGSVTVYCGGSLSPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 240
QY 233 LHKDYSADTLAHNNDIALLKIRSKGRCQAPSRITQICLPSMYNDPQFGTSCEITGFGK 292
DB 241 LHKDYSADTLAHNNDIALLKIRSKGRCQAPSRITQICLPSMYNDPQFGTSCEITGFGK 300
QY 293 ENSTDYLYPEQLKMTVVKLIHSHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGPL 352
DB 301 ENSTDYLYPEQLKMTVVKLIHSHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGPL 360
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 403
DB 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

RESULT 10
US-08-560-098A-47
; Sequence 47. Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEEDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/560.098A
 FILING DATE: 17-NOV-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: P 44 40 892.7
 FILING DATE: 17-NOV-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: EVANS, Joseph D.
 REGISTRATION NUMBER: 26,269
 REFERENCE/DOCKET NUMBER: 148/42448
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-8800
 TELEFAX: (202) 628-8844
 INFORMATION FOR SEQ ID NO: 47:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 432 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-560-098A-47

Query Match 87.1%; Score 1965; DB 2; Length 432;
 Best Local Similarity 90.5%; Pred. No. 4.5e-164;
 Matches 361; Conservative 8; Mismatches 16; Indels 14; Gaps 2;
 QY 13 CLNGTCVSNKYFNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
 DB 40 CVTGEPTPKPSHNNGDPEEIPPEY-----LQISKTCYEGNGHFYRGKASTDTMGRPCL 93
 QY 73 PWSNATVLLQOYTHAHSRSDALQGLGKHNCRPNRRPWCYVQVGLKPLVQECMVHDC 132
 DB 94 PWSNATVLLQOYTHAHSRSDALQGLGKHNCRPNRRPWCYVQVGLKPLVQECMVHDC 153
 QY 133 DCK-----LKFCGGQKTLRPRKLIIGGFTIENQPFPAALYRRHGGSVTYVCGGS 184
 DB 154 DKKPSSPPEELKFCGGQKTLRPRKLIIGGFTIENQPFPAALYRRHGGSVTYVCGGS 213
 QY 185 LISPCWVISATHCFIDYPKKEDIYVLRGRSLNSNTQEMKFEVENILHKDYSADTLAH 244
 DB 214 LISPCWVISATHCFIDYPKKEDIYVLRGRSLNSNTQEMKFEVENILHKDYSADTLAH 273
 QY 245 HNDIALKIRSKGRCACPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQL 304
 DB 274 HNDIALKIRSKGRCACPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQL 333
 QY 305 KMTVVKLIASHRECQPHYVGGSEVTKMLCAADPQWKTDSCGDSGGPLVCSLQGRMTLTG 364
 DB 334 KMTVVKLIASHRECQPHYVGGSEVTKMLCAADPQWKTDSCGDSGGPLVCSLQGRMTLTG 393
 QY 365 IVSWGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
 DB 394 IVSWGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 432

RESULT 11
 US-08-093-741-83
 ; Sequence 83, Application US/08093741
 ; Patent No. 5681721
 ; GENERAL INFORMATION:
 ; APPLICANT: STEFFENS, GERD J.
 ; APPLICANT: WNEEDT, STEPHAN
 ; APPLICANT: SCHNEIDER, JOHANNES
 ; APPLICANT: HEINZEL-WIELAND, REGINA
 ; APPLICANT: SAUNDERS, DEREK J.
 ; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
 ; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
 ; TITLE OF INVENTION: INHIBITING EFFECT
 ; NUMBER OF SEQUENCES: 83
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
 STREET: 1200 G Street, N. W. Suite 700
 CITY: Washington, D.C.
 COUNTRY: U.S.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/093,741
 FILING DATE: 20-JUL-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE P43 23 754.1
 FILING DATE: 15-JUL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: EVANS, Joseph D.
 REGISTRATION NUMBER: 26,269
 REFERENCE/DOCKET NUMBER: 148/41345
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-8800
 TELEFAX: (202) 628-8844
 INFORMATION FOR SEQ ID NO: 83:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 365 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-093-741-83

Query Match 87.0%; Score 1964; DB 1; Length 365;
 Best Local Similarity 97.8%; Pred. No. 4.4e-164;
 Matches 357; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWSNATVLLQOYTHAHSRSDALQGLGKHNCRNP 106
 DB 1 SKTCYEGNGHFYRGKASTDTMGRPCLPWSNATVLLQOYTHAHSRSDALQGLGKHNCRNP 60
 QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGK-----LKFQCGQKTLRPRFKIIGGFTT 158
 DB 61 NRRRPWCYVQVGLKPLVQECMVHDCADGKPPSPPEELKFCGGQKTLRPRFKIIGGFTT 120
 QY 159 IENQPFPAALYRRHGGSVTYVCGSLISPCWVISATHCFIDYPKKEDIYVLRGRSLNS 218
 DB 121 IENQPFPAALYRRHGGSVTYVCGSLISPCWVISATHCFIDYPKKEDIYVLRGRSLNS 180
 QY 219 NTQGMKFEVENILHKDYSADTLAHNDIALKIRSKGRCACPSRTIQTICLPSMYND 278
 DB 181 NTQGMKFEVENILHKDYSADTLAHNDIALKIRSKGRCACPSRTIQTICLPSMYND 240
 QY 279 PQFGTSCEITGFGKENSTDYLYPEQLKMTVVKLIASHRECQPHYVGGSEVTKMLCAADPQ 338
 DB 241 PQFGTSCEITGFGKENSTDYLYPEQLKMTVVKLIASHRECQPHYVGGSEVTKMLCAADPQ 300
 QY 339 WKTDSCGDSGGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKEE 398
 DB 301 WKTDSCGDSGGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKEE 360
 QY 399 NGLAL 403
 DB 361 NGLAL 365

RESULT 12
 US-08-720-012-83
 ; Sequence 83, Application US/08720012
 ; Patent No. 5747291
 ; GENERAL INFORMATION:
 ; APPLICANT: STEFFENS, GERD J.
 ; APPLICANT: WNEEDT, STEPHAN
 ; APPLICANT: SCHNEIDER, JOHANNES
 ; APPLICANT: HEINZEL-WIELAND, REGINA

APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
TITLE OF INVENTION: INHIBITING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N. W. Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/720,012
FILING DATE: 27-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/093,741
FILING DATE: 20-JUL-1993
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-720-012-83

Query Match 87.0%; Score 1964; DB 1; Length 365;
Best Local Similarity 97.8%; Pred. No. 4.4e-164;
Matches 357; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPD 106
DB 1 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPD 60
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTT 158
DB 61 NRRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEELKFQCGQKTLRPRFKIIGGEFTT 120
QY 159 IENQPFAAIYRRHRGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRRLNS 218
DB 121 IENQPFAAIYRRHRGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRRLNS 180
QY 219 NTQGEKMFVENLILHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPMSYND 278
DB 181 NTQGEKMFVENLILHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPMSYND 240
QY 279 PQFGTSCETGFGKENS TDVLYPEQLKMTVVKLISHRECQOPHYGSEVTTKMLCAADPQ 338
DB 241 PQFGTSCETGFGKENS TDVLYPEQLKMTVVKLISHRECQOPHYGSEVTTKMLCAADPQ 300
QY 339 WKTDS CGDGGPLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKEE 398
DB 301 WKTDS CGDGGPLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKEE 360
QY 399 NGLAL 403
DB 361 NGLAL 365

RESULT 13

US-08-560-098A-44
Sequence 44, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENNDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having Fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-44

Query Match 87.0%; Score 1964; DB 2; Length 393;
Best Local Similarity 97.8%; Pred. No. 4.9e-164;
Matches 357; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPD 106
DB 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPD 61
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTT 158
DB 62 NRRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEELKFQCGQKTLRPRFKIIGGEFTT 121
QY 159 IENQPFAAIYRRHRGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRRLNS 218
DB 122 IENQPFAAIYRRHRGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRRLNS 181
QY 219 NTQGEKMFVENLILHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPMSYND 278
DB 182 NTQGEKMFVENLILHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPMSYND 241
QY 279 PQFGTSCETGFGKENS TDVLYPEQLKMTVVKLISHRECQOPHYGSEVTTKMLCAADPQ 338
DB 242 PQFGTSCETGFGKENS TDVLYPEQLKMTVVKLISHRECQOPHYGSEVTTKMLCAADPQ 301
QY 339 WKTDS CGDGGPLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKEE 398
DB 302 WKTDS CGDGGPLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKEE 361
QY 399 NGLAL 403

Db 362 NGLAL 366

RESULT 14

US-08-967-024C-24

; Sequence 24, Application US/08967024C

; Patent No. 6133011

; GENERAL INFORMATION:

; APPLICANT: WNEBTD, Stephan

; APPLICANT: STEFFENS, Gerd Josef

; APPLICANT: JANOSHA, Elke

; APPLICANT: HEINZEL-WIELAND, Regina

; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan

; STREET: 1200 G Street, N.W., Suite 700

; CITY: Washington

; STATE: DC

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/967,024C

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: P 44 42 665.8

; FILING DATE: 30-NOV-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: EVANS, Joseph D.

; REGISTRATION NUMBER: 26,269

; REFERENCE/DOCKET NUMBER: 148/42444

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 628-8800

; TELEFAX: (202) 628-8844

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 393 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-967-024C-24

Query Match 87.0%; Score 1964; DB 3; Length 393;

Best Local Similarity 97.8%; Pred. No. 4.9e-164;

Matches 357; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 47 SKTCYEGNGHFRGKASTDTMGRCPLPNSATVLCQTYHAHRSALQLGLGKHNCRND 106

Db 2 SKTCYEGNGHFRGKASTDTMGRCPLPNSATVLCQTYHAHRSALQLGLGKHNCRND 61

Qy 107 NRRPWCYVQVGLKPLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTT 158

Db 62 NRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEELKFCGQKTLRPRFKIIGGEFTT 121

Qy 159 IENQPFWFAAIYRRHRGGSVTVCCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNS 218

Db 122 IENQPFWFAAIYRRHRGGSVTVCCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNS 181

Qy 219 NTQGENKFEVENLILHKDYSADTLAHNDIALLLKRSKEGRCAQPSRTIQTICLPMSYND 278

Db 242 POFGTSCEITGFGKENSNDYLYPEQLKMTVVKLISHRECQPHYYGSEVITKMLCAADPQ 301

Qy 339 WKTDSCQDGGPLVCSLQGRMTLTGIYWGRCALKDKPGVYTRVSHFLPWIRSHTKEE 398

Db 302 WKTDSCQDGGPLVCSLQGRMTLTGIYWGRCALKDKPGVYTRVSHFLPWIRSHTKEE 361

Qy 399 NGLAL 403

Db 362 NGLAL 366

RESULT 15

US-08-967-024C-25

; Sequence 25, Application US/08967024C

; Patent No. 6133011

; GENERAL INFORMATION:

; APPLICANT: WNEBTD, Stephan

; APPLICANT: STEFFENS, Gerd Josef

; APPLICANT: JANOSHA, Elke

; APPLICANT: HEINZEL-WIELAND, Regina

; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan

; STREET: 1200 G Street, N.W., Suite 700

; CITY: Washington

; STATE: DC

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/967,024C

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: P 44 42 665.8

; FILING DATE: 30-NOV-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: EVANS, Joseph D.

; REGISTRATION NUMBER: 26,269

; REFERENCE/DOCKET NUMBER: 148/42444

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 628-8800

; TELEFAX: (202) 628-8844

; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 393 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-967-024C-25

Query Match 87.0%; Score 1964; DB 3; Length 393;

Best Local Similarity 97.8%; Pred. No. 4.9e-164;

Matches 357; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 47 SKTCYEGNGHFRGKASTDTMGRCPLPNSATVLCQTYHAHRSALQLGLGKHNCRND 106

Db 2 SKTCYEGNGHFRGKASTDTMGRCPLPNSATVLCQTYHAHRSALQLGLGKHNCRND 61

Qy 107 NRRPWCYVQVGLKPLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTT 158

Db 62 NRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEELKFCGQKTLRPRFKIIGGEFTT 121

Qy 159 IENQPFWFAAIYRRHRGGSVTVCCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNS 218

Db 122 IENQPFWFAAIYRRHRGGSVTVCCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNS 181

Qy 219 NTQGENKFEVENLILHKDYSADTLAHNDIALLLKRSKEGRCAQPSRTIQTICLPMSYND 278

Db 182 NTQGENKFEVENLILHKDYSADTLAHNDIALLLKRSKEGRCAQPSRTIQTICLPMSYND 241

QY 279 POFGTSCEITGKENSVDYLYPEQLKMTVVVKLISHRECCQPHYVYGVSEVTTKMLCAADPQ 338
DB 242 POFGTSCEITGKENSVDYLYPEQLKMTVVVKLISHRECCQPHYVYGVSEVTTKMLCAADPQ 301
QY 339 WKTSQCGSDSGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEE 398
DB 302 WKTSQCGSDSGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEE 361
QY 399 NGLAL 403
DB 362 NGLAL 366

RESULT 16
US-08-560-098A-45
; Sequence 45, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-45

Query Match 65.0%; Score 1466.5; DB 2; Length 306;
Best Local Similarity 91.1%; Pred. No. 1.3e-120;
Matches 275; Conservative 6; Mismatches 12; Indels 9; Gaps 3;

QY 103 RNPDRRRP-WCYVQVGLKPLVQECMVHDCADGKLFQCGOKTLRPRFKIIGGEFTTIIEN 161
DB 13 RNPNDKYEFWEDEEKG--PHMSS-----PPEELKFCQCGOKTLRPRFKIIGGEFTTIIEN 64
QY 162 QPWFAAIYRRHRGGSVTVYCGGSLISPCWVVSATHCFIDYPKKEDYIVYLGSRSLNNTQ 221
DB 65 QPWFAAIYRRHRGGSVTVYCGGSLISPCWVVSATHCFIDYPKKEDYIVYLGSRSLNNTQ 124
QY 222 GEMKFEVENLILHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPSMYNDPOF 281

DB 125 GEMKFEVENLILHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPSMYNDPOF 184
QY 282 GTSCEITGKENSVDYLYPEQLKMTVVVKLISHRECCQPHYVYGVSEVTTKMLCAADPQWK 341
DB 185 GTSCEITGKENSVDYLYPEQLKMTVVVKLISHRECCQPHYVYGVSEVTTKMLCAADPQWK 244
QY 342 DSCQSDSGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGL 401
DB 245 DSCQSDSGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGL 304
QY 402 AL 403
DB 305 AL 306

RESULT 17
US-08-560-098A-46
; Sequence 46, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-46

Query Match 65.0%; Score 1466.5; DB 2; Length 331;
Best Local Similarity 91.1%; Pred. No. 1.4e-120;
Matches 275; Conservative 6; Mismatches 12; Indels 9; Gaps 3;

QY 103 RNPDRRRP-WCYVQVGLKPLVQECMVHDCADGKLFQCGOKTLRPRFKIIGGEFTTIIEN 161
DB 13 RNPNDKYEFWEDEEKG--PHMSS-----PPEELKFCQCGOKTLRPRFKIIGGEFTTIIEN 64
QY 162 QPWFAAIYRRHRGGSVTVYCGGSLISPCWVVSATHCFIDYPKKEDYIVYLGSRSLNNTQ 221
DB 65 QPWFAAIYRRHRGGSVTVYCGGSLISPCWVVSATHCFIDYPKKEDYIVYLGSRSLNNTQ 124
QY 222 GEMKFEVENLILHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLPSMYNDPOF 281

Db 125 GEMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMTNDFQF 184
QY 282 GTSCEITGFGKENSVDLYPEQLKMTVWKLISHRECQPHYGVSEVTTKMLCAADPQWKT 341
Db 185 GTSCEITGFGKENSVDLYPEQLKMTVWKLISHRECQPHYGVSEVTTKMLCAADPQWKT 244
QY 342 DSCQDGGGGLVCSLQGRMTLTGIVSWGRCALKKPKGVYTRVSHFLPWIRSHTKENGL 401
Db 245 DSCQDGGGGLVCSLQGRMTLTGIVSWGRCALKKPKGVYTRVSHFLPWIRSHTKENGL 304
QY 402 AL 403
Db 305 AL 306

RESULT 18
US-08-944-483-73
; Sequence 73, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; TITLE OF INVENTION: OF THE PROSTATE
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: linear
US-08-944-483-73
Query Match 61.2%; Score 1382; DB 3; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.5e-113;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 IIGGEFTTIENQPFAAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPPKEDYIV 210
Db 1 IIGGEFTTIENQPFAAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPPKEDYIV 60
QY 211 LGRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQT 270
Db 61 LGRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQT 120
QY 271 CLPSMYNDPQFGTSCEITGFGKENSVDLYPEQLKMTVWKLISHRECQPHYGVSEVTTK 330
Db 121 CLPSMYNDPQFGTSCEITGFGKENSVDLYPEQLKMTVWKLISHRECQPHYGVSEVTTK 180
QY 331 MLCAADPQWKTDCQDGGGLVCSLQGRMTLTGIVSWGRCALKKPKGVYTRVSHFLPW 390
Db 181 MLCAADPQWKTDCQDGGGLVCSLQGRMTLTGIVSWGRCALKKPKGVYTRVSHFLPW 240
QY 391 IRSHTKEENGLAL 403
Db 241 IRSHTKEENGLAL 253

RESULT 19
US-08-560-098A-49
; Sequence 49, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WEINZEL, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-49
Query Match 60.9%; Score 1374; DB 2; Length 254;
Best Local Similarity 99.2%; Pred. No. 1.3e-112;
Matches 252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 150 KIIGGEFTTIENQPFAAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPPKEDYIV 209
Db 1 KIIGGEFTTIENQPFAAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPPKEDYIV 60

QY 210 YLGRSLNSNTQGMKFEVENLILHKDYADTLAHNDIALLKIRSKGRCQAQPSRTIQT 269
DB 61 YLGRSLNSNTQGMKFEVENLILHKDYADTLAHNDIALLKIRSKGRCQAQPSRTIQT 120
QY 270 ICLPSMYNDPQFGTSCEITGFGKENSTDLYPEQLKMTVVKLIHRECCQPHYGVSEVTT 329
DB 121 ICLPSMYNDPQFGTSCEITGFGKENSTDLYPEQLKMTVVKLIHRECCQPHYGVSEVTT 180
QY 330 KMLCAADPQWKTDSQCGSGPLVCSLQGRMTLGTIVSWGRGCALKDKPGVYTRVSHFLP 389
DB 181 KMLCAADPQWKTDSQCGSGPLVCSLQGRMTLGTIVSWGRGCALKDKPGVYTRVSHFLP 240
QY 330 WIRSHTEENGLAL 403
DB 241 WIRSHTEENGLVL 254

RESULT 20

US-09-101-272G-73
; Sequence 73, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: ATF domain of uPA
US-09-101-272G-73

Query Match 43.6%; Score 983; DB 4; Length 200;
Best Local Similarity 95.6%; Pred. No. 1.6e-78;
Matches 172; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHYRG 60
DB 21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHYRG 80
QY 61 KASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPDNRRPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPDNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQGKTLPRPKIIGGEFTTIENQPFPAIYRRH 172
DB 141 PLVQECMVHDCADGKSPSPPEELKFCQGKTLPRPKIIGGEFTTIENQPFPAIYRRH 200

RESULT 21

US-09-560-098A-51
; Sequence 51 Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEZTL, Stephan
; APPLICANT: HEINZEL, Wieland, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evensen, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington

; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-51

Query Match 38.4%; Score 867.5; DB 2; Length 477;
Best Local Similarity 43.7%; Pred. No. 6.1e-68;
Matches 178; Conservative 60; Mismatches 148; Indels 21; Gaps 9;

QY 3 ELHQVP-----SNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHY 58
DB 78 QCHTVFVKSCSELRCFCNGGTCQAASFSDF-VQCPKGYTGKQCEVDTHATCYKQGVY 136
QY 59 RKASTDITMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPDNRRPWCYVQV 118
DB 137 RGTWSTSEGAQCINWNSNLLTRTYNGRRSDAITLGLGNHNYCRNPDNRRPWCYVKA 196
QY 119 LKPLVQECMVHDCADGKLFQCG-QKTLPRPKIIGGEFTTIENQPFPAIYRRHGGG- 176
DB 197 SKFLFCVPSVCS-----KATGLRKRYKEPQLHSTGGLFTDITSHFWAAIFQNRSSG 252
QY 177 VTYVCGGSLISPCWISATHCFID-YPKKEDIYVILGRSLNSNTQGMKFEVENLILHK 235
DB 253 ERFLCGGILISSCWLVTAACFQERYPPQHLAVV-LGRTYRVKPGKEEQTFEVEKCIVHE 311
QY 236 DYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTCLPSMYNDPQFGTSCEITGFGKENS 295
DB 312 EFDDDT--YNDIALQLKSGSPQCAQESDSVRAICLPEANLQLPDWTECELSGYGKHKS 369
QY 296 TDLYPEQLKMTVVKLIHRECCQPHYGVSEVTTKMLCAADPQWKT-----DSCQDGS 349
DB 370 SSPFFYSEQLKEGHVRLYPSRCTSKFLFNKTVTKMLCAGDTRSGEIHFNVDACQDGS 429
QY 350 GPLVCSLQGRMTLGTIVSWGRGCALKDKPGVYTRVSHFLPWRSHTK 396
DB 430 GPLVCRNDNHNMTLLGIISWVGCGEKDIPGVYTKVTNVLGWIRDNR 476

RESULT 22

US-08-811-949-39
; Sequence 39, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI

```

; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO. 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-811-949-39

```

RESULT 24
US-07-609-510B-16
; Sequence 16, Application US/07609510B
; Patent No. 5326700
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Method for Al-

```

; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN.
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,510B
; FILING DATE: 19901106
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-609-510B-16

Query Match 38.0%; Score 858.5; DB 1; Length 527;
Best Local Similarity 37.5%; Pred. No. 4.3e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

Qy 3 ELHQVP-SNCD---CLNGGTCVSNKYFNTHWCNPKKFGGQHCIDKSKTCYEGNGHY 58
Db 42 QCHSVPVKSCSPRCFNGGTCQALYFSDF-VQCPEGFAGKCEIDTRATCYEDQGISY 100
Qy 59 RGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNDRRPPWCYVQVG 118
Db 101 RGTWSTAESGAECTNWNSSALAQKPYSGRRPDAIRLGLGNHNYCRNPNDRSDKPCYVFK 160
Qy 119 LKPLVQECMVHDCADG-----KLKFG-----CG-QKTLRPRFKIIGGE 155
Db 161 GYISSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWNMSMILIGKVYTAQNP 220
Qy 135 -----KLKFG-----CG-QKTLRPRFKIIGGE 155
Db 221 AQALGLGKHNYCRNPDGDAKPMCHLVKNRRLTWECYDVPSCSTCGLRQYQPFRIKGL 280
Qy 156 FTTIENQWFAAIYRRH-RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVVLGRS 214
Db 281 FADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGR 340
Qy 215 RLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRSKGRCACQPSRTIOTICLPS 274
Db 341 YRVVPGEEQKEVEKYIVHKEFDDDT--YNDIALQLKSDSSRCAQESSVVRIVCLPP 398
Qy 275 MYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCA 334
Db 399 ADLQLPDWTECELSGKGHEALSPYSERLKEAHVLYPSSRCTSQHLLNRTVTDNMLCA 458
Qy 335 AD-----PQWKT-DSCQDGGGLVCSLQGRMILTGIVSWGRGCAKDKPGVYTVSHFL 388
Db 459 GDRSGGPQANLHDACQDGGGLVCLNDGRMTLVGIISWGLGCGQKDVFGVYTKVTNYL 518
Qy 389 PWIRSHTK 396
Db 519 DWIRDNMR 526

RESULT 25
PCT-US91-01025A-2
; Sequence 2, Application PC/TUS9101025A
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin

```

```

; TITLE OF INVENTION: Specific Properties
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01025A
; FILING DATE: 19910214
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/486,657
; FILING DATE: 1 March 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasek, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 454P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; PCT-US91-01025A-2

```

```

Query Match 38.0%; Score 858.5; DB 5; Length 527;
Best Local Similarity 37.5%; Pred. No. 4.3e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

Qy 3 ELHQVP-SNCD---CLNGGTCVSNKYFNTHWCNPKKFGGQHCIDKSKTCYEGNGHY 58
Db 42 QCHSVPVKSCSPRCFNGGTCQALYFSDF-VQCPEGFAGKCEIDTRATCYEDQGISY 100
Qy 59 RGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNDRRPPWCYVQVG 118
Db 101 RGTWSTAESGAECTNWNSSALAQKPYSGRRPDAIRLGLGNHNYCRNPNDRSDKPCYVFK 160
Qy 119 LKPLVQECMVHDCADG-----KLKFG-----CG-QKTLRPRFKIIGGE 155
Db 161 GYISSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWNMSMILIGKVYTAQNP 220
Qy 135 -----KLKFG-----CG-QKTLRPRFKIIGGE 155
Db 221 AQALGLGKHNYCRNPDGDAKPMCHLVKNRRLTWECYDVPSCSTCGLRQYQPFRIKGL 280
Qy 156 FTTIENQWFAAIYRRH-RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVVLGRS 214
Db 281 FADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGR 340
Qy 215 RLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRSKGRCACQPSRTIOTICLPS 274
Db 341 YRVVPGEEQKEVEKYIVHKEFDDDT--YNDIALQLKSDSSRCAQESSVVRIVCLPP 398
Qy 275 MYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCA 334
Db 399 ADLQLPDWTECELSGKGHEALSPYSERLKEAHVLYPSSRCTSQHLLNRTVTDNMLCA 458
Qy 335 AD-----PQWKT-DSCQDGGGLVCSLQGRMILTGIVSWGRGCAKDKPGVYTVSHFL 388
Db 459 GDRSGGPQANLHDACQDGGGLVCLNDGRMTLVGIISWGLGCGQKDVFGVYTKVTNYL 518
Qy 389 PWIRSHTK 396

```


335 AD-----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRGKALKDKPGVTVRSHFL 388
 494 GTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGLIISWGLGCGQKQDVPVTVTKVTNYL 553
 QY 389 PWIRSHTK 396
 Db 554 DWIRDNR 561
 RESULT 28
 US-08-560-098A-50
 ; Sequence 50, Application US/08560098A
 ; Patent No. 5976841
 ; GENERAL INFORMATION:
 ; APPLICANT: WENNDT, Stephan
 ; APPLICANT: HEINZEL-WIELAND, Regina
 ; APPLICANT: STEFFENS, Gerd Josef
 ; TITLE OF INVENTION: Proteins having Fibrinolytic and
 ; TITLE OF INVENTION: Coagulation-inhibiting Properties
 ; NUMBER OF SEQUENCES: 60
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
 ; STREET: 1200 G Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/560,098A
 ; FILING DATE: 17-NOV-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: P 44 40 892.7
 ; FILING DATE: 17-NOV-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: EVANS, Joseph D.
 ; REGISTRATION NUMBER: 26,269
 ; REFERENCE/DOCKET NUMBER: 148/42448
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 628-8800
 ; TELEFAX: (202) 628-8844
 ; INFORMATION FOR SEQ ID NO: 50:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 562 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-560-098A-50
 Query Match 38.0%; Score 858.5; DB 2; Length 562;
 Best Local Similarity 37.5%; Pred. No. 4.6e-67;
 Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
 QY 3 ELHQP-SNCD---CLNGGTCVSNKYFENHWCNCPKFGGCHCEIDKSKTCYEGNGHFY 58
 Db 77 QCHSVFVKSCEPRFCNGTQQALYFSDP-VQCPEGFAGKCCEDTRATCYEDQGISY 135
 QY 59 RKGASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRRCWYQVG 118
 Db 136 RGTWSTAESAECTNWNSSALAQKPYSGRPPDAIRLGLGNHNYCRNPNDRDSDKWCYVFK 195
 QY 119 LKELVQECMVDHDCADG-----
 Db 196 GKYSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWN SMILIGKVYTAQNES 255
 QY 135 -----KLKFKQ-----CG-QKTLRPRFKLIGGE 155
 Db 256 AQALGLGKHNYCRNPDGADKPWCHVLKNRRLTWEYCDVPSCSTGLRQYSQPQFRKGG 315

QY 156 FTIENQWFAAIYRRH-RGGSVTVCCGSLSPCMWISATCFIDYPKKEDIYVYLGSR 214
 Db 316 FADIASHWQAAIFAFHARRSPGERFLCGGILISSCWILSAHCFQBERPPPHLTVILGRT 375
 QY 215 RLNSNTQGMKEFVENLILHKDYSADTLAHNNDIALKLRKSGRCQAPSRITQITCLPS 274
 Db 376 YRVVPEEEQKEVEKIVHKEFDDDT--YNDNIALQLKSDSSRCAQESSVVRTVCLPP 433
 QY 275 MYNDPQFGTSCBITGFKENSTDIYLPQLKMTVVKLISHRECQPHYVYGVSEVTMLCA 334
 Db 434 ADLQLPDWTECBLSGYKHEALSPFYSERLKEAHLVLPSSRCTSOHLNLRVTIONMLCA 493
 QY 335 AD-----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRGKALKDKPGVTVRSHFL 388
 Db 494 GTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGLIISWGLGCGQKQDVPVTVTKVTNYL 553
 QY 389 PWIRSHTK 396
 Db 554 DWIRDNR 561
 RESULT 29
 US-08-883-795A-38
 ; Sequence 38, Application US/08883795A
 ; Patent No. 5985607
 ; GENERAL INFORMATION:
 ; APPLICANT: Delcuve, Genevieve
 ; APPLICANT: Awang, Gregor
 ; TITLE OF INVENTION: Recombinant DNA Molecules and Expression
 ; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
 ; NUMBER OF SEQUENCES: 39
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BERESKIN & PARR
 ; STREET: 40 King Street West
 ; CITY: Toronto
 ; STATE: Ontario
 ; COUNTRY: Canada
 ; ZIP: M5H 3Y2
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/883,795A
 ; FILING DATE: 27-JUN-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gravelle, Micheline
 ; REGISTRATION NUMBER: 40,261
 ; REFERENCE/DOCKET NUMBER: 7841-062
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (416) 364-7311
 ; TELEFAX: (416) 361-1398
 ; INFORMATION FOR SEQ ID NO: 38:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 562 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein (tpa)
 ; US-08-883-795A-38
 Query Match 38.0%; Score 858.5; DB 2; Length 562;
 Best Local Similarity 37.5%; Pred. No. 4.6e-67;
 Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
 QY 3 ELHQP-SNCD---CLNGGTCVSNKYFENHWCNCPKFGGCHCEIDKSKTCYEGNGHFY 58
 Db 77 QCHSVFVKSCEPRFCNGTQQALYFSDP-VQCPEGFAGKCCEDTRATCYEDQGISY 135
 QY 59 RKGASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRRCWYQVG 118


```
Db 136 RGTWSTAEGACETNWNSSALAQPSYGRPPDAIRLGLGNHNYCRNPDROKPCWYVKA 195
Qy 119 LKPLVQECMVHDCADG-----
Db 196 GKYSEFCSTPACSEGNDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKYVTAQNP 255
Qy 135 -----KLKFO-----CG-OXTLRPRFKIIG 155
Db 256 AQALGLGKHNYCRNPDGAKPWCHLVKNRRLTWECVDPSCSTCGLRQYSPQFRIKGL 315
Qy 156 FTTIENQFWFAAIYRRH-RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGRS 214
Db 316 FADIASHPWQAAIAFAKHRRSPGERFLOGGILISSCWILSSAAHCFQERPPPHLTVILGR 375
Qy 215 RLNSNTQGEKMFVENILHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIOTICLPS 274
Db 376 YRVVPEEEQKFEVEKYIVHKEFDDT--YNDIALQLKSDSSRCAQESSVVRTVCLPP 433
Qy 275 MYNDPQFCTSCITGFGKENSTDYLYPEOLKMTVVKLISHRECOCPHYGSEVTTKMLCA 334
Db 434 ADLQLPDWTCELSYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRVTDNMLCA 493
Qy 335 AD-----PWKMT-DSCQDSGGPLVCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFL 388
Db 494 GDTSGGPQANLHDACQDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVFGVYTKVTNYL 553
Qy 389 PWIRSHTK 396
Db 554 DWIRDNMR 561

RESULT 30
5185259-3
; Patent No. 5185259
; APPLICANT: GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; VEHR, GORDON A.
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; SEQ ID NO:3
; LENGTH: 562
5185259-3

Query Match 38.0%; Score 858.5; DB 6; Length 562;
Best Local Similarity 37.5%; Pred. No. 4.6e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

Qy 3 BLHQP-SNCD---CLNGGTCVSNKYFSNIHWCNPKKFGQHCIEDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSEPRCFNGGTCQALYFSDF-VCQCPGCFAGKCEIDTRATCYEDQGISY 135
Qy 59 RGKASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGUGKHNYCRNPDNRPRWCYVQVG 118
Db 136 RGTWSTAEGACETNWNSSALAQPSYGRPPDAIRLGLGNHNYCRNPDROKPCWYVKA 195
Qy 119 LKPLVQECMVHDCADG-----
Db 196 GKYSEFCSTPACSEGNDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKYVTAQNP 255
Qy 135 -----KLKFO-----CG-OXTLRPRFKIIG 155
```

```
Db 256 AQALGLGKHNYCRNPDGAKPWCHLVKNRRLTWECVDPSCSTCGLRQYSPQFRIKGL 315
Qy 156 FTTIENQFWFAAIYRRH-RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGRS 214
Db 316 FADIASHPWQAAIAFAKHRRSPGERFLOGGILISSCWILSSAAHCFQERPPPHLTVILGR 375
Qy 215 RLNSNTQGEKMFVENILHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIOTICLPS 274
Db 376 YRVVPEEEQKFEVEKYIVHKEFDDT--YNDIALQLKSDSSRCAQESSVVRTVCLPP 433
Qy 275 MYNDPQFCTSCITGFGKENSTDYLYPEOLKMTVVKLISHRECOCPHYGSEVTTKMLCA 334
Db 434 ADLQLPDWTCELSYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRVTDNMLCA 493
Qy 335 AD-----PWKMT-DSCQDSGGPLVCSLQGRMTLTGIVSWGRGKALKDKPGVYTRVSHFL 388
Db 494 GDTSGGPQANLHDACQDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVFGVYTKVTNYL 553
Qy 389 PWIRSHTK 396
Db 554 DWIRDNMR 561
```

Search completed: December 3, 2003, 14:45:37
Job time : 22.8448 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:58 ; Search time 20.8448 Seconds
(without alignments)
1859.261 Million cell updates/sec

Title: US-09-880-503-6

Perfect score: 2257

Sequence: 1 SNELHQVPSNCCLNGGTCV.....VSHFLPWRSHKTEENGLAL 403

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2240	99.2	431	1 UKHU	u-plasminogen acti
2	2090	92.6	433	1 UKBAY	u-plasminogen acti
3	1827.5	81.0	442	1 UKPG	u-plasminogen acti
4	1728	76.6	433	1 JN0560	u-plasminogen acti
5	1656.5	73.4	432	1 S18932	u-plasminogen acti
6	1626.5	72.1	433	1 UKMS	u-plasminogen acti
7	997	44.2	434	1 A35005	u-plasminogen acti
8	868.5	38.5	477	2 JS0598	t-plasminogen acti
9	857.5	38.4	477	1 A34369	t-plasminogen acti
10	852.5	38.2	431	2 JS0599	t-plasminogen acti
11	858.5	38.0	562	1 UKHUT	t-plasminogen acti
12	841.5	37.3	559	1 A35039	t-plasminogen acti
13	840.5	37.2	477	2 JS0597	t-plasminogen acti
14	829.5	36.8	559	1 A29941	t-plasminogen acti
15	756	33.5	394	2 JS0600	t-plasminogen acti
16	726.5	32.2	655	1 A46688	hepatocyte growth
17	700.5	31.0	603	2 S28941	coagulation factor
18	681	30.2	615	1 KPHU12	coagulation factor
19	664.5	29.4	558	2 JC5878	plasma hyaluronan-
20	651.5	28.9	560	1 JC4795	plasma hyaluronan-
21	638	28.3	593	2 S45281	coagulation factor
22	508.5	22.5	790	1 PLPG	plasmin (EC 3.4.21
23	502	22.2	460	2 B61545	plasmin (EC 3.4.21
24	501	22.2	810	1 PLHU	plasmin (EC 3.4.21
25	501	22.2	810	2 B30848	plasmin (EC 3.4.21
26	501	22.2	812	1 PLMS	plasmin (EC 3.4.21
27	484.5	21.5	812	1 PLBO	plasmin (EC 3.4.21
28	478	21.2	4548	1 S00657	apoptoprotein(a) (EC
29	477	21.1	1420	2 A32869	apoptoprotein(a)

ALIGNMENTS

RESULT 1

UKHU

u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N;Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen
N;Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen a
in form
C;Species: Homo sapiens (man)
C;Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text change 15-Sep-2000
C;Accession: A00931; I52209; JTO102; A37561; J38102; S65783; A37562; A37564; A;
R;Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Biasi, P.
Nucleic Acids Res. 13, 2759-2771, 1985
A;Title: The human urokinase-plasminogen activator gene and its promoter.
A;Reference number: A00931; MUID:85215647; PMID:2987867
A;Accession: A00931
A;Molecule type: DNA
A;Residues: 1-431 <PIC>
A;Cross-references: GB:X02419; NID:937601; PIDN:CAA26268.1; PID:G1834524
A;Note: the authors translated the codon ATG for residue 214 as Ile
R;Nagamine, Y.; Pearson, D.; Grattan, M.
Biochem. Biophys. Res. Commun. 132, 563-569, 1985
A;Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcin
A;Reference number: I52209; MUID:86050639; PMID:3933505
A;Accession: I52209
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: DNA
A;Residues: 145-161 <NAG1>
A;Cross-references: GB:K03027; NID:G340174; PIDN:AAA61257.1; PID:G340175
R;Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama,
Gene 36, 183-188, 1985
A;Title: Molecular cloning of cDNA coding for human preprourokinase.
A;Reference number: JTO102; MUID:86056954; PMID:2415429
A;Accession: JTO102
A;Molecule type: mRNA
A;Residues: 1-213, 'I', 215-431 <NAG2>
A;Cross-references: GB:K03226; NID:G340155; PIDN:AA097138.1; PID:G340158; GB:D00244; N;
R;Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Biasi, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984
A;Title: Identification and primary sequence of an unspliced human urokinase poly(A) + r
A;Reference number: A37561; MUID:84272706; PMID:6589620
A;Accession: A37561
A;Molecule type: mRNA
A;Residues: 66-431 <VER>
A;Cross-references: GB:D00244; NID:G220138
R;Jacobs, P.; Cravador, A.; Lortau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elser
DNA 4, 139-146, 1985
A;Title: Molecular cloning, sequencing, and expression in Escherichia coli of human pre
A;Reference number: I38102; MUID:85203359; PMID:3888571
A;Accession: I38102
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>
A;Cross-references: EMBL:X02760; NID:G35297; PIDN:CAA26535.1; PID:G35298

30 451.5 20.0 455 2 A61545 plasmin (EC 3.4.21
31 434 19.2 761 2 JC5759 brain-specific ser
32 431 19.1 810 2 I46260 plasmin (EC 3.4.21
33 408.5 18.1 343 1 A57014 proctasin (EC 3.4.
34 407.5 18.1 638 1 KQMSPL plasma kallikrein
35 404.5 17.9 417 1 S00845 hepsin (EC 3.4.21.
36 402 17.8 638 1 KQHUP plasma kallikrein
37 401 17.8 416 1 S33777 hepsin (EC 3.4.21.
38 397 17.6 855 2 JC7731 membrane-bound arg
39 395.5 17.5 248 2 S55066 trypsin (EC 3.4.21
40 394.5 17.5 229 1 TRBOTR trypsin (EC 3.4.21
41 390.5 17.3 247 2 S19813 trypsin (EC 3.4.21
42 390 17.3 263 2 A21195 chymotrypsin (EC 3
43 387.5 17.2 407 1 KPSO7 coagulation factor
44 386.5 17.1 638 1 KQRTPL plasma kallikrein
45 383 17.0 461 1 KXHU protein C (activat

R.Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, Biochim. Biophys. Acta 1293, 83-89, 1996
 A>Title: Characterization of single chain urokinase-type plasminogen activator with a no
 A:Reference number: S65783; MUID:96186279; PMID:8652631
 A:Accession: S65783
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 21-140; 'L', 142-213; 'I', 215-431 <YOS>
 A:CROSS-references: ENBLD11143; NID:91311467; PIDN:BA001919.1; PID:g1199928
 R:Gunzler, W.A.; Steffens, G.J.; Oetting, F.; Kim, S.M.A.; Frankus, E.; Flohe, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
 A>Title: The primary structure of high molecular mass urokinase from human urine.
 A:Reference number: A37562; MUID:83055084; PMID:6754569
 A:Accession: A37562
 A:Molecule type: protein
 A:Residues: 21-177; <GUN>
 R:Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O. Eur. J. Biochem. 125, 251-257, 1982
 A>Title: Human low-molecular-weight urinary urokinase. Partial characterization and pre
 A:Reference number: A37563; MUID:83003608; PMID:6749491
 A:Accession: A37563
 A:Molecule type: protein
 A:Residues: 156-176; 179-193; 'T', 195; 'T', 197-224 <SCH>
 R:Steffens, G.J.; Gunzler, W.A.; Oetting, F.; Frankus, E.; Flohe, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
 A>Title: The complete amino acid sequence of low molecular mass urokinase from human uri
 A:Reference number: A37564; MUID:83055099; PMID:6754572
 A:Accession: A37564
 A:Molecule type: protein
 A:Residues: 158-410 <STE>
 R:Kentzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K. Biochem. Biophys. Res. Commun. 171, 401-406, 1990
 A>Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinan
 A:Reference number: A35689; MUID:90365737; PMID:2393398
 A:Accession: A35689
 A:Molecule type: protein
 A:Residues: 21-30; 'X', 32; 'X', 34-38; 'X', 40-43 <KEN>
 A>Note: identification of a fucose and attempt to determine its attachment site
 R:Rabani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltzm Biochim. Biophys. Res. Commun. 173, 1058-1064, 1990
 A>Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li
 A:Reference number: A36697; MUID:91097529; PMID:2125213
 A:Accession: A36697
 A:Molecule type: protein
 A:Residues: 21-34 <RAB>
 R:Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M. submitted to the Brookhaven Protein Data Bank, July 1993
 A:Reference number: A51255; PDB:1KDU
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R:Li, X.; Smith, R.A.G.; Dobson, C.M. Biochemistry 31, 9562-9571, 1992
 A>Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain b
 A:Reference number: A44375; MUID:93003110; PMID:1327118
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
 R:Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazar, A.P.; Olejniczak, submitted to the Brookhaven Protein Data Bank, January 1994
 A:Reference number: A68822; PDB:1URK
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R:Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.; submitted to the Brookhaven Protein Data Bank, July 1995
 A:Reference number: A56038; PDB:1LW
 A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426
 C:Comment: This enzyme is found in urine in a high molecular mass form, consisting of A
 C:Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a
 C:Genetics:
 A:Gene: GDB:PLAU
 A:CROSS-references: GDB:119497; OMIM:191840
 A:Map position: 10q24-10q24
 A:Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3
 C:Function:
 A:Description: proteolytically activates plasminogen
 A:Pathway: fibrinolysis
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C;Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine protease
 F;1-20/Domain: signal sequence #status predicted <SIG>
 F;21-431/Product: urokinase-type plasminogen activator, single chain form #status predi
 F;21-177/Product: urokinase-type plasminogen activator chain A #status experimental <ME
 F;31-62/Domain: EGF homology <EGF>
 F;70-151/Domain: kringle homology <KRG>
 F;156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental <
 F;179-431/Product: urokinase-type plasminogen activator chain B #status experimental <B
 F;179-419/Domain: trypsin homology <TRY>
 F;31-39,33-53,53-62,70-151,91-133,122-146,168-299,209-225,217-288,313-382,345-361,372-4
 F;38/Binding site: carbohydrate (Thr) (covalent) #status predicted
 F;178-179/Cleavage site: Lys-Ile (plasmin) #status experimental
 F;224,275,376/Active site: His, Asp, Ser #status experimental
 F;322/Binding site: carbohydrate (Asn) (covalent) #status experimental
 Query Match 99.2%; Score 2240; DB 1; Length 431;
 Best Local Similarity 97.8%; Pred. No. 2,3e-166;
 Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;
 QY 1 SNELHQPNSCCLNGTCTVSNKYFSNIHWCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
 DB 21 SNELHQPNSCCLNGTCTVSNKYFSNIHWCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGRCPLPWSNATVLQOTYHAHRSDALQLGLGKHNCRPNRRRPPWCYVOVGUK 120
 DB 81 KASTDTMGRCPLPWSNATVLQOTYHAHRSDALQLGLGKHNCRPNRRRPPWCYVOVGUK 140
 QY 121 PLVQECMVHDCADGK-----LKFCQCKTLRPRFKIIGGEFTTIENQPFATYYRSH 172
 DB 141 PLVQECMVHDCADGKPPSPPELKFQCCQKTLRPRFKIIGGEFTTIENQPFATYYRSH 200
 QY 173 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLI 232
 DB 201 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLI 260
 QY 233 LHKQYSADTLAHNDIALLKIRSKRCACQAPRTTCTICLPMSYNDPQGTSCETIGFGK 292
 DB 261 LHKQYSADTLAHNDIALLKIRSKRCACQAPRTTCTICLPMSYNDPQGTSCETIGFGK 320
 QY 293 ENSTDLYPEQLKMTYVVKLISHRECOQPHYVSEVTTKMLCAADPOWKTDSQCGSGGGL 352
 DB 321 ENSTDLYPEQLKMTYVVKLISHRECOQPHYVSEVTTKMLCAADPOWKTDSQCGSGGGL 380
 QY 353 VCSLQGRMTLTGIVSWGRGCAKDKPFGVYTRVSHFLPWIRSHTKENGUAL 403
 DB 381 VCSLQGRMTLTGIVSWGRGCAKDKPFGVYTRVSHFLPWIRSHTKENGUAL 431
 RESULT 2
 UKBAY
 u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
 C:Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
 C:Accession: S14687; S08651
 R:Au, Y.P.T.; Wang, T.W.; Clowes, A.W. Nucleic Acids Res. 18, 3411, 1990
 A>Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminog
 A:Reference number: S14687; MUID:90287734; PMID:2113276
 A:Accession: S14687
 A:Molecule type: mRNA
 A:Residues: 1-433 <AU>
 A:CROSS-references: EMBL:X51935; NID:G38130; PIDN:CAA36200.1; PID:g38131
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
 C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F;1-20/Domain: signal sequence #status predicted <SIG>
 F;21-176/Product: plasminogen activator chain A #status predicted <ACH>
 F;30-61/Domain: EGF homology <EGF>
 F;69-150/Domain: kringle homology <KRG>
 F;178-433/Product: plasminogen activator chain B #status predicted <SCH>
 F;178-421/Domain: trypsin homology <TRY>
 F;167-298,208-224,216-287,315-384,347-363,374-403/Disulfide bonds: #status predicted
 F;223,274,376/Active site: His, Asp, Ser #status predicted
 F;324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 92.6%; Score 2090; DB 1; Length 433;
Best Local Similarity 90.6%; Pred. No. 1e-154;
Matches 375; Conservative 17; Mismatches 10; Indels 12; Gaps 3;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 21 SREL-QVPSDCGCLNGGTCMSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 79

QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 80 KASDTMTGRSCLANWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 139

QY 121 PLVQECMVHDCADGK-----LKFCQCGQKTLRPRFKIIGGEFTTIENQFWFAAIYRRH 172
DB 140 QRVQECMVHDCADGKPPSEELQFCQCGQKTLRPRFKIIGGEFTTIENQFWFAAIYRRH 199

QY 173 RGGSVYVCGGSLISPCWVWISATHCFIDYPKKEDYIVYLGSRNSNTQGEKFEVENLI 232
DB 200 RGGSVYVCGGSLISPCWVWISATHCFINYPKKEDYIVYLGSRNSNTQGEKFEVENLI 259

QY 233 LHKDYSADTLAHHNDIALLLKIRSEGRCAQPSRTIQTICLPSMYNDPQ---FGTSCBITG 289
DB 260 LHEDYSADTLAHHNDIALLLKIRSEGRCAQPSRTIQTICLPSMYNDPQFPFGTSCBITG 319

QY 290 FGKNSDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTSCQDGS 349
DB 320 FGKNSDYLYPEQLKMTVVKLVSHQKQCPHYGVSEVTTKMLCAADPQWKTSCQDGS 379

QY 350 GPLVCSLQGRMTLTGIVSWGRCALKKPKGVYTVRSVHFLPWIRSHTKENGLAL 403
DB 380 GPLVCSLQGRMTLTGIVSWGRCALKKPKGVYTVRSVHFLPWIRSHTKENGLAL 433

RESULT 3
UKPG
u-plasminogen activator (EC 3.4.21.73) precursor - pig
N:Alternate names: uPA
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
A:Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A:Reference number: A00932; MUID:85087954; PMID:6096832
A:Accession: A00932
A:Molecule type: DNA
A:Residues: 1-240; 'H', 242-442 <NAG1>
A:Experimental source: kidney cell line LLC-PK1
R:Nagamine, Y.
submitted to the Protein Sequence Database, December 1986
A:Reference number: A37566
A:Contents: annotation; correction to residue 241
C:Genetics:
A:Insertions: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
F:152/Binding site: carbohydrate (Asn)
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-189/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:33-64/Domain: EGF homology <EGF>
F:72-153/Domain: kringle homology <KR>
F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:190-430/Domain: trypsin homology <TRY>
F:152/Binding site: carbohydrate (Asn)
F:179-310,220-236,228-259,324-393,356-372,383-411/Disulfide bonds: #status predicted
F:235,286,387/Active site: His, Asp, Ser #status predicted

Query Match 81.0%; Score 1827.5; DB 1; Length 442;
Best Local Similarity 78.1%; Pred. No. 2.5e-134;
Matches 329; Conservative 32; Mismatches 41; Indels 19; Gaps 2;

QY 1 SNEHQV--PSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFY 58

DB 21 SHELHQESGASNCGLNGKCKVSYKFSNIQRSCPKKFGQBCEIDTSQTCFEGNGHSY 80

QY 59 RGAASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQV 118

DB 81 RGAASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQV 140

QY 119 LKPLVQECMVHDCADG-----KLKFCQCGQKTLRPRFKIIGGEFTTIEN 161

DB 141 LKQVQECMVHDCADG-----KLKFCQCGQKTLRPRFKIIGGEFTTIEN 200

QY 162 QPFWFAIYRRHGGSVYVCGGSLISPCWVWISATHCFIDYPKKEDYIVYLGSRNSNTQ 221

DB 201 QPFWFAIYRRHGGSVYVCGGSLISPCWVWISATHCFINYPKKEDYIVYLGSRNSNTQ 250

QY 222 GEMKFEVENLIHKDYSADTLAHHNDIALLLKIRSEGRCAQPSRTIQTICLPSMYNDPQ 281

DB 261 GEMKFEVENLIHKDYSADTLAHHNDIALLLKIRSEGRCAQPSRTIQTICLPSMYNDPQ 320

QY 282 GTSCEITGKNSDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKT 341

DB 321 GASCEITGKNSDYLYPEQLKMTVVKLVSHRECOQPHYGVSEVTTKMLCAADPQWKT 380

QY 342 DSCQDGSGLVCSLQGRMTLTGIVSWGRCALKKPKGVYTVRSVHFLPWIRSHTKENGL 401

DB 381 DSCQDGSGLVCSLQGRMTLTGIVSWGRCALKKPKGVYTVRSVHFLPWIRSHTKENGL 440

QY 402 A 402

DB 441 A 441

RESULT 4
JN0560
u-plasminogen activator (EC 3.4.21.73) precursor - bovine
N:Alternate names: uPA
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C:Accession: JN0560
R:Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
Gene 125, 177-183, 1993
A:Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and indi
A:Reference number: JN0560; MUID:93216119; PMID:8385052
A:Accession: JN0560
A:Molecule type: mRNA
A:Residues: 1-433 <KRA>
C:Cross-references: GB:L03546; NID:g163800; PID:AAA51419.1; PID:g163801
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-179/Product: plasminogen activator chain A #status predicted <MA1>
F:21-179/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:33-64/Domain: EGF homology <EGF>
F:72-153/Domain: kringle homology <KRG>
F:181-433/Product: plasminogen activator chain B #status predicted <MA2>
F:181-421/Domain: trypsin homology <TRY>
F:170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
F:226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 76.6%; Score 1728; DB 1; Length 433;
Best Local Similarity 73.8%; Pred. No. 1.3e-126;
Matches 305; Conservative 44; Mismatches 54; Indels 10; Gaps 2;

QY 1 SNEHQV--PSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFY 58

DB 21 SNEVKHSESGSNCGLNGKCKVSYKFSNIQRSCPKKFGQBCEIDTSQTCFEGNGHSY 80

QY 59 RGAASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQV 118

DB 81 RGAASDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQV 140

QY 119 LKPLVQECMVHDCADGK-----KFCQCGQKTLRPRFKIIGGEFTTIENQFWFAAIYR 170

DB 141 LKQVQECMVHDCADGK-----KFCQCGQKTLRPRFKIIGGEFTTIENQFWFAAIYR 200

```

QY 171 RRRGSSVTVCCGSLISPCWVISAATHCFIDYPKKEDYIVYLGSRSLNNTQGMKFEVEN 230
D 171 RRRGSSVTVCCGSLISPCWVISAATHCFIDYPKKEDYIVYLGSRSLNNTQGMKFEVEN 230
D 201 RRRGSSVTVCCGSLISPCWVISAATHCFIDYPKKEDYIVYLGSRSLNNTQGMKFEVEN 260
QY 231 LILHKDYSADTLAHNDIALKIRSKGRCAQPSRTIOTICLPSMYNDPQFGTSCIEITGF 290
D 231 LILHKDYSADTLAHNDIALKIRSKGRCAQPSRTIOTICLPSMYNDPQFGTSCIEITGF 290
D 261 LILHEDYSAESLAHNDIALKIRSKGRCAQPSRTIOTICLPSMYNDPQFGTSCIEITGF 320
QY 291 GKENSTDYLPQLKMTVVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCGDSGG 350
D 291 GKENSTDYLPQLKMTVVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCGDSGG 350
D 321 GKENSTDYLPQLKMTVVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCGDSGG 380
QY 351 PLVCSLQGRMTITGVISWGRGKALDKPGVYTRVSHFPLWIRSHTKKEENGLA 403
D 351 PLVCSLQGRMTITGVISWGRGKALDKPGVYTRVSHFPLWIRSHTKKEENGLA 403
D 381 PLVCTIQGLTITGVISWGRGKALDKPGVYTRVSHFPLWIRSHTKKEENGLA 433

RESULT 5
S18932
u-plasminogen activator (EC 3.4.21.73) precursor - rat
N:Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator
C:Species: Rattus norvegicus (Norway rat)
C>Date: 18-Oct-1989 #sequence revision 10-Feb-1995 #text_change 18-Jun-1999
C:Accession: S24604; I60186; I53472; S18932
R:Kabbani, S.A.
submitted to the EMBL Data Library, April 1992
A:Reference number: S24604
A:Accession: S24604
A:Molecule type: mRNA
A:Residues: 1-15, 'H', 17-23, 'G', 25-331, 'N', 333-432 <RAB>
A:Cross-references: EMBL:X65851; NID:G57456; PIDN:CAA46601.1; PID:G57457
A:Experimental source: tissue kidney
R:Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.
Cancer Res. 52, 2489-2496, 1992
A:Title: Transcriptional and posttranscriptional activation of urokinase plasminogen act
A:Reference number: I60186; MUID:92233409; PMID:1568219
A:Accession: I60186
A>Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-432 <RES>
A:Cross-references: EMBL:X63434; NID:957465; PIDN:CAA45028.1; PID:G57466
A:Experimental source: strain Fischer 344; tissue mammary
R:Ragno, P.; Cassano, S.; Degen, J.; Kessler, C.; Blasi, F.; Rossi, G.
FEBS Lett. 306, 193-198, 1992
A:Title: The receptor for the plasminogen activator of urokinase type is up-regulated in
A:Reference number: I53472; MUID:92339549; PMID:1321734
A:Accession: I53472
A>Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 31-62 <RE2>
A:Cross-references: EMBL:X66907; NID:G396200; PIDN:CAA47356.1; PID:G938279
C:Genetics:
A:Gene: uPA
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:31-62/Domain: EGF homology <EGF>
F:70-151/Domain: kringle homology <KRG>
F:179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:179-420/Domain: trypsin homology <TRY>
F:168-300, 210-226, 218-289, 314-383, 346-362, 373-401/Disulfide bonds: #status predicted
F:225, 276, 377/Active site: His, Asp, Ser #status predicted

Query Match 73.4%; Score 1656.5; DB 1; Length 432;
Best Local Similarity 72.2%; Pred No. 4.7e-12;
Matches 291; Conservative 45; Mismatches 50; Indels 9; Gaps 2;

QY 9 SNCDCLNGGTCVSNKYFNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFRGHASTDTMG 68
D 9 SNCDCLNGGTCVSNKYFNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFRGHASTDTMG 68
D 29 SNCGCQNGGVCVSYKFSIRRCSCPKFGQEHCEIDKSKTCYHGNGDSYRGKANTDTKG 88

RESULT 6
UKMS
u-plasminogen activator (EC 3.4.21.73) precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1987 #sequence revision 30-Sep-1987 #text_change 18-Jun-1999
C:Accession: A29420; A24615
R:Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
Biochemistry 26, 8270-8279, 1987
A:Title: The murine urokinase-type plasminogen activator gene.
A:Reference number: A29420; MUID:88163489; PMID:2831940
A:Accession: A29420
A:Molecule type: DNA
A:Residues: 1-433 <DEG>
A:Cross-references: GB:M17922; NID:G202296; PIDN:AAA40539.1; PID:G202297
R:Belin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Koche
Eur. J. Biochem. 148, 225-232, 1985
A:Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase
A:Reference number: A24615
A:Accession: A24615
A:Molecule type: mRNA
A:Residues: 1-433 <BEL>
A:Cross-references: GB:X02389; NID:G55127; PIDN:CAA26231.1; PID:G55128
C:Genetics:
A:Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:32-63/Domain: EGF homology <EGF>
F:71-152/Domain: kringle homology <KRG>
F:180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:180-421/Domain: trypsin homology <TRY>
F:129-301, 211-227, 219-230, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
F:226, 277, 378/Active site: His, Asp, Ser #status predicted

Query Match 72.1%; Score 1626.5; DB 1; Length 433;
Best Local Similarity 70.2%; Pred. No. 1e-118;
Matches 283; Conservative 50; Mismatches 61; Indels 9; Gaps 2;

QY 9 SNCDCLNGGTCVSNKYFNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFRGHASTDTMG 68
D 9 SNCDCLNGGTCVSNKYFNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFRGHASTDTMG 68
D 30 SNCGCQNGGVCVSYKFSIRRCSCPKFGQEHCEIDKSKTCYHGNGDSYRGKANTDTKG 89
QY 69 RPCLPWNSATVLOQTYHAHSDALQGLGKHNYCRPNDRRRPWCYVQVGLKPLVQECMV 128
D 69 RPCLPWNSATVLOQTYHAHSDALQGLGKHNYCRPNDRRRPWCYVQVGLKPLVQECMV 128
D 90 RPCLANNAVLOKPNNAHNPDAISLGLGKHNYCRPNDRPWCYVQVGLKPLVQECMV 149
QY 129 HCDADGKLK-----FQCGQKTLRPRFKIIGGFTTINQWPFAAIYRHRGGS-VTY 179
D 129 HCDADGKLK-----FQCGQKTLRPRFKIIGGFTTINQWPFAAIYRHRGGS-VTY 179

```

Db 150 HDCSLSKPSSVDPQGGCCQCKALRPRFKIVGGEFTVEVENQWFAAIYQKNKGSSPPSF 209
Qy 180 VCGSLISPCWVWISATHCFIYPKKEDYIVYLGSRSLNSNTQGMKEFEVENLILHKDYSA 239
Db 210 KCGSLISPCWVVAACHCFIOLPKENVYVLOGSKSSYNPGMKFEVQLLIHYYRE 269
Qy 240 DTLAHHNDIALKIRSKREGCAQPSRTIQTICLPSMYNDPQFGTSCEITFGKENSTDYL 299
Db 270 DSLAYHNDIALKIRSTGQCAQPSRTIQTICLPRFPDAPFGSDCEITFGKESSESYL 329
Qy 300 YPEOLKMTVKVLLSHRECOQPHHYGSEVTTKMLCAADPQWKTDCQDSCGDSGLVCSLQGR 359
Db 330 YPKMLKMSVVKLVSHRECOQPHHYGSEVTTKMLCAADPQWKTDCQDSCGDSGLVCSLQGR 369
Qy 360 MTLTGIVSWRGCAKDKPGVYTRVSHFLPWIRSHTEENGLA 402
Db 390 PTLTGIVSWRGCAKDKPGVYTRVSHFLPWIRSHTEENGLA 432
RESULT 7
A:Species: Gallus gallus (chicken)
C:Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 16-Jul-1999
C:Accession: A35005
R:Leslie, N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.
J. Biol. Chem. 265, 1339-1344, 1990
A:Title: The chicken urokinase-type plasminogen activator gene.
A:Reference number: A35005; MUID:90110185; PMID:2295632
A:Accession: A35005
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-434 <LS>
A:Cross-references: GB:J05187; NID:G212858; PID:AAA49131.1; PID:G212859
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:40-71/Domain: EGF homology <EGF>
F:79-158/Domain: kringle homology <KRG>
F:173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:173-416/Domain: trypsin homology <TRY>
F:162-296,202-218,210-285,310-379,342-358,369-397/Disulfide bonds: #status predicted
F:217,272,373/Active site: His, Asp, Ser #status predicted
Query Match 44.2%; Score 997; DB 1; Length 434;
Best Local Similarity 47.8%; Pred. No. 7.5e-70;
Matches 187; Conservative 64; Mismatches 122; Indels 18; Gaps 7;
Qy 11 CDCINGTCVSNKYFSGNIHWCNCPKFGGQHCIDSKTCYEGNGHFYRGKASTDMGRP 70
Db 40 CQCLNGTCITYRFFSOIKELCPGEGVGHCHCIDTNSICYNGEDYRQMAEDP----G 95
Qy 71 CLPWSATVLO-QTYHAHRSDALQLGLGKXNYCRNPNRRPWCYVQVGLKPLVQF--CM 127
Db 96 CLYWDHPSVIRWGDYHADLKNALQLGLGKXNYCRNPNRRPWCYTK--RRYSIQETPCS 153
Qy 128 VHCADGKLKFCQGGKTLRPRFKIIGFEFTIENQWFAAIYRRHGGSVTYVCGSLIS 187
Db 154 TIE---KCERTCGQSFYSFYKIVGSGQAEVETQWIAIFQNIK-GTDQFLCGGSLID 208
Qy 188 PCWVISATHCFID---YPKKEDYIVYLGSRSLNSNTQGMKEFEVENLILHKDYSAADTLA 243
Db 209 PCWVLTAAHCFYNPTKQPKNSYKVFGLGKSLTNDHEQVFWDEIISHPFDTHDTCG 268
Qy 244 HNNDIALKIRSKREGCAQPSRTIQTICLPSMYNDPQFGTSCEITFGKENSTDYLYPEQ 303
Db 269 NNDIALIRIRASGQCAVESNTRVTCPLKPNLNYDNTWCEIAGYQKNSYDIYYAQR 328
Qy 304 LKMTVKVLLSHRECOQPHHYGSEVTTKMLCAADPQWKTDCQDSCGDSGLVCSLQGRMTLT 363

Db 329 LMSATVNLISQDDCKNKYYDSTRVTDNMVCAGDPLMETDACKDSCGPMVCEHNGRMTLY 388
Qy 364 GIVSWGRCGALKDKPGVYTRVSHFLPWIRSH 394
Db 389 GIVSWGRCGALKDKPGVYTRVSHFLPWIRSH 419
RESULT 8
JS0598
t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0598
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; De
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0598
A:Molecule type: mRNA
A:Residues: 1-477 <KRA>
A:Cross-references: GB:M63988; NID:G166074; PID:AAA11593.1; PID:G166075
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hc
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <IFA>
F:128-209/Domain: EGF homology <EGF>
F:226-471/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-
F:185,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
F:272,321,428/Active site: His, Asp, Ser #status predicted
Query Match 38.5%; Score 868.5; DB 2; Length 477;
Best Local Similarity 43.7%; Pred. No. 7.8e-60;
Matches 178; Conservative 60; Mismatches 148; Indels 21; Gaps 9;
Qy 3 ELHQVP----SNCDCLNGTCVSNKYFSGNIHWCNCPKFGGQHCIDSKTCYEGNGHFY 58
Db 78 QCHTVPVKSCSELRCFNGGTCWQAASFSDF-VCCQPKGYTGKQCEVDTHATCYKQGVTY 136
Qy 59 RKASTDTMGRLCPWNSATVQQTYHAHRSDALQLGLGKXNYCRNPNRRPWCYVQVQ 118
Db 137 RGTWSTSEGAQCINWNSNLLTRTYNGRRSDAITLGLGNHNYCRNPNNSKPCVYTKA 196
Qy 119 LKPLVQECMVHDCADGKLKFCQG-OKTLRPRFKIIGFEFTIENQWFAAIYRRHGGSV- 176
Db 197 SKFIEFCVPCVCS---KATCGLRKYKEPQHSTGGGLTDTITSHPQAAIIPAQRSSG 252
Qy 177 VTYVCGGSLISPCWVISATHCFID-YPKKEDYIVYLGSRSLNSNTQGMKEFEVENLILHK 235
Db 253 ERFLLCGGILLSCWVLTAAHCFQERYPPQHLRW-LGRTYVYKPGKEQTTEVEKCIYHE 311
Qy 236 DYSADTLAHHNDIALKIRSKREGCAQPSRTIQTICLPSMYNDPQFGTSCEITFGKENS 295
Db 312 EFDDDT--YNNDIALLQKSGSPQCAQESDSVRAICLPEANLQLPDMTECELSGYGKHS 369
Qy 236 TDLYPEOLKMTVKVLLSHRECOQPHHYGSEVTTKMLCAADPQWKT-----DSCQDGG 349
Db 370 SSPFFYSEQLKEGHVLYPSSRCTSKFLNKVTNNMLCAGDTRSGEYIPNVHDCQDGG 429
Qy 350 GPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTK 396
Db 430 GPLVCMNDNMHTLLGIISWGVGCGEKDIPGVYTKVNYLGIWRDMR 476
RESULT 9
A34369
t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
C:Species: Megaderma lyra

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C/Accession: A34369
 R/Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacob
 J. Biol. Chem. 264, 17947-17952, 1989
 A>Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmin
 A/Reference number: A34369; MUID:90036867; PMID:2509450
 A/Accession: A34369
 A>Status: Preliminary
 A/Molecule type: mRNA
 A/Residues: 1-477 <GAR>
 A/Cross-references: GB:J05082; NID:G166080; PIDN:AAA31596.1; PID:G166081
 C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-36/Domain: propeptide #status predicted <PRO>
 F:37-477/Product: plasminogen activator #status predicted <PLA>
 F:42-79/Domain: fibronectin type I repeat homology <1FA>
 F:87-120/Domain: EGF homology <EGF>
 F:128-209/Domain: kringle homology <KRG>
 F:226-471/Domain: trypsin homology <TRY>
 F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
 F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 38.4%; Score 867.5; DB 1; Length 477;
 Best Local Similarity 43.7%; Pred. No. 9.4e-60;
 Matches 178; Conservative 60; Mismatches 148; Indels 21; Gaps 9;

QY 3 ELHQP-----SNCDLNGGTCVSNKYFSNIHWCNPKKFGGHCIDKSKTCYEGNGHFY 58
 DB 78 QCHTVPVKSCSELRCFNGGTCWQAASFDF-VCQCPKGYTGKQCEVDTHATCYKDQGVY 136
 QY 59 RKASTDTRGRCPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRRPPCYQVQG 118
 DB 137 RGTWSTSEGAQCINWNSNLLTRTYNGRRSDAITGLGNHNYCRNPNRRRPPCYQVYKA 196
 QY 119 LKPLVQECWHDGKLFQCG-QKTLRPRFKIIGGEFTTIENQWFAAIYRRHRGGS- 176
 DB 197 SKFLFCFVPVCS-----KATCGLRKYKPEQLHSTGGLFTDITSHPWQAIFAQNRSSG 252
 QY 177 VTYVCGSLISPCWVISATHCFID-YPKKEDIYVYLGSRSLNSNTQGEKFEVENLILHK 235
 DB 253 ERFGLGILISSCWLTAAHCFQERYPPQHLRVV-LGRTYRVKPKKEEQTEVEKCIHVE 311
 QY 236 DYSADTLAHNDIALKRSKEGRCAPQSRITQICLPSMYNDPQGTSCITGFGKENS 295
 DB 312 EFDDDT--YNNDIALLQKSGSPQCAQESDSVRAICLPEANLQLPDWTCELSGYGRHS 369
 QY 296 TDLYPEQLKMTVVVLIHRECCQPHYGYSEVTTKMLCAADPQWKT-----DSCQGDG 349
 DB 370 SSPFYSEQLKGVHLYPSSRCTSKFLFNKVTNNMLCAGTRSGEHPNVHDAQCQDGS 429
 QY 350 GPLVCSLOGRWTLTGIVSWGRGCKALDKPGVYTRVSHFLPWIRSHTK 396
 DB 430 GPLVCRNDNHTLLGIISWGVGCGEKDIPGVYTKVTNVLGWIRDNMR 476

RESULT 10
 JS0599
 t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
 N/Alternate names: tissue plasminogen activator
 C/Species: Desmodus rotundus (common vampire bat)
 C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C/Accession: JS0599
 R/Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don
 Gene 105, 229-237, 1991
 A>Title: The plasminogen activator family from the salivary gland of the vampire bat Des
 A/Reference number: JS0599; MUID:92039036; PMID:1937019
 A/Accession: JS0599
 A/Molecule type: mRNA
 A/Residues: 1-431 <KRA>
 A/Cross-references: GB:M63989; NID:G166076; PIDN:AAA31594.1; PID:G166077
 C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase

F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-36/Domain: propeptide #status predicted <PRO>
 F:37-431/Product: plasminogen activator beta #status predicted <PLA>
 F:41-74/Domain: EGF homology <EGF>
 F:82-163/Domain: kringle homology <KRG>
 F:180-425/Domain: trypsin homology <TRY>
 F:41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disulfide bo
 F:139,352/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:179-180/Cleavage site: His-Ser (plasmin) #status predicted
 F:226,275,382/Active site: His, Asp, Ser #status predicted
 F:345-361,378-406/Disulfide bonds: #status predicted

Query Match 38.2%; Score 862.5; DB 2; Length 431;
 Best Local Similarity 44.1%; Pred. No. 2e-59;
 Matches 175; Conservative 59; Mismatches 146; Indels 17; Gaps 8;

QY 9 SNCDLNGGTCVSNKYFSNIHWCNPKKFGGHCIDKSKTCYEGNGHFYRGKASTDTMG 68
 DB 42 SELRCFNGGTCWQAASFDF-VCQCPKGYTGKQCEVDTHATCYKDQGVYRGWSTSESG 100
 QY 69 RPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRRPPCYQVGLKPLVQECMV 128
 DB 101 AQCNWNSNLLTRTYNGRRSDAITGLGNHNYCRNPNRRRPPCYQVYKASKFILEFCV 160
 QY 129 HCDAGKLFQCG-QKTLRPRFKIIGGEFTTIENQWFAAIYRRHRGGS-VTYVCGGSLI 186
 DB 161 PVCS-----KATCGLRKYKPEQLHSTGGLFTDITSHPWQAIFAQNRSSGFLCGILI 216
 QY 187 SPQWVISATHCFID-YPKKEDIYVYLGSRSLNSNTQGEKFEVENLILHKVSAADTLAH 245
 DB 217 SSCWLTAAHCFQERYPPQHLRVV-LGRTYRVKPKKEEQTEVEKCIHIEEFDOT--YN 273
 QY 246 NDIALKRSKEGRCAPQSRITQICLPSMYNDPQGTSCITGFGKENSVDLYPEQLK 305
 DB 274 NDIALQKSGSPQCAQESDSVRAICLPEANLQLPDWTCELSGYGRHSKSSPFYSEQLK 333
 QY 306 MTYVLIHRECCQPHYGYSEVTTKMLCAADPQWKT-----DSCQGDGGLVCSLOGR 359
 DB 334 EGHVLYPSSRCTSKFLFNKVTNNMLCAGTRSGEHPNVHDAQCQDGGPLVCMNDNH 393
 QY 360 MTLTGIVSWGRGCKALDKPGVYTRVSHFLPWIRSHTK 396
 DB 394 MTLGLIISWGVGCGEKDIPGVYTKVTNVLGWIRDNMR 430

RESULT 11
 UKHUT
 t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
 N/Alternate names: t-PA; tissue plasminogen activator
 C/Species: Homo sapiens (man)
 C/Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 08-Dec-2000
 C/Accession: A94004; A23529; J0562; A93293; S02125; A91343; A93951; A91322; A54645; I
 R/NY, T.; Elgh, F.; Lund, B.
 Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
 A>Title: The structure of the human tissue-type plasminogen activator gene: correlation
 A/Reference number: A94004; MUID:84298137; PMID:6089198
 A/Accession: A94004
 A/Molecule type: DNA
 A/Residues: 1-562 <NYT>
 A/Cross-references: GB:I00141
 A/Note: the codon given for residue 93 (ACC) is inconsistent with the authors' transla
 R/Priener Degen, S.J.; Rajpur, B.; Reich, E.
 J. Biol. Chem. 261, 6972-6985, 1986
 A>Title: The human tissue plasminogen activator gene.
 A/Reference number: A23529; MUID:86196143; PMID:3009482
 A/Accession: A23529
 A/Molecule type: DNA
 A/Residues: 1-562 <DEG>
 A/Cross-references: GB:K03021; NID:G339817; PIDN:AAA98809.1; PID:G339818
 R/Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.
 Agric. Biol. Chem. 55, 1225-1232, 1991
 A>Title: Purification and characterization of tissue plasminogen activator secreted by
 A/Reference number: J0562; MUID:91291340; PMID:1368681

```

Query Match      38.0%; Score 858.5; DB 1; Length 562;
Best Local Similarity 37.5%; Pred. No. 5.6e-59;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

QY   3 ELHQVP-SNCD---CLNGTCTVSNKVFPSNIHWNCNPKPKFGQGHCEIDSKSTCYEGNHGFY 58
    :|::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db   77 QCHSVTPKSCBERCFNGTCQQALYFSDP-VQCPEGFAGKCCCIDTRATCYEDQGISY 135
    :|::|||:|||||:|||||:|||||:|||||:|||||:|||||:
QY   59 RGKASDTWTGRPCLPWN SATVILOQTTHAURSDALQLGLGKHNYCRNPDRRRPMWCYYQVG 118
    :|::|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db   136 RGTWSTAESGAECTNNWSALLAQKPYSGRPPDAIRLGLGNHNYCRNPDRDSKPKWCYFEKA 195
    :|::|||:|||||:|||||:|||||:|||||:|||||:|||||:
QY   119 LXLPLVECMVHCADG----- 134
    :|::|||:|||||:|||||:|||||:|||||:|||||:
Db   196 GKYSSEFCSTPACSEGNSDCYFNGSAYRGTHTSLTESGASCLPWNSMLIGKVYTAQNPS 255
    :|::|||:|||||:|||||:|||||:|||||:|||||:
QY   135 -----KLKEFO-----CG-OXTLRPRFKITGE 155
    :|::|||:|||||:|||||:|||||:|||||:|||||:
Db   256 AQAALGLGKHNYCRNP DGDAPKWPCHVLKNRLTW EYCDVPFSCSTCGLRQYSPQFRIRKGL 315
    :|::|||:|||||:|||||:|||||:|||||:|||||:
QY   156 FTTIENQPFAAIYRRH-RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDIVYLGRS 214
    :|::|||:|||||:|||||:|||||:|||||:|||||:

```


Db 316 FADIASHPQAAIFAKHRSRPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRT 375
QY 215 RLNSNTQGMKEPEVENLILHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIOTICLPS 274
Db 376 YRVVPGEEOKEFEKIVHKEFDODT--YNDIALLOKSDSRCAQESSWRTVCLPP 433
QY 275 MYNDPQFCTSCITGKENSVDLYPEOLKMTVVKLISHRECCOPHYGSEVTTKMLCA 334
Db 434 ADLQLPDWTCELSYGKHEALSPYSERLKEAHVRLYPSRCSQHLNRTVTDNMLCA 493
QY 335 AD-----PQWKT-DSQCDSGGPLVCSLQGRMTLTGIVSWRGCAKDKPGVYTRVSHFL 388
Db 494 GDRSGGQANLHDACQDSGGPLVCLNDGRMTLVGIIISWGLGCGQKDPGVGKVTNYL 553
QY 389 PWIRSHTK 396
Db 554 DWIRDNMR 561
RESULT 12
t-plasminogen activator (EC 3.4.21.68) precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A35029; A31597
R:Feng, P.; Ohlsson, M.; Ny, T.
J. Biol. Chem. 265, 2022-2027, 1990
A:Title: The structure of the TATA-less rat tissue-type plasminogen activator gene. Spec
A:Reference number: A35029; MUID:90130448; PMID:2105315
A:Accession: A35029
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-559 <FEN>
A:Cross-references: GB:M31197; NID:G207429; PIDN:AAA42261.1; PID:G207431; GB:J05226
R:NY, T.; Leonardsson, G.; Haueh, A.J.W.
DNA 7, 671-677, 1988
A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activator
A:Reference number: A31597; MUID:89170114; PMID:3148445
A:Accession: A31597
A:Molecule type: mRNA
A:Residues: 1-379, 'K', 381-559 <NYT>
A:Cross-references: GB:M33697; NID:G530159; PIDN:AAA41812.1; PID:G530160
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-29/Domain: propeptide #status predicted <PRO>
F:30-559/Product: t-plasminogen activator #status predicted <MAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F:38-75/Domain: fibronectin type I repeat homology <1F1>
F:83-116/Domain: EGF homology <EGF>
F:124-205/Domain: kringle homology <KR1>
F:213-294/Domain: kringle homology <KR2>
F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F:309-553/Domain: trypsin homology <TRY>
F:38-68, 66-75, 93-94, 88-105, 107-116, 124-205, 145-187, 176-200, 213-294, 234-276, 265-289, 297-4
F:149, 481/Binding site: carboxylate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F:355, 404, 510/Active site: His, Asp, Ser #status predicted
Query Match 37.3%; Score 841.5; DB 1; Length 559;
Best Local Similarity 36.5%; Pred. No. 1.2e-57;
Matches 180; Conservative 65; Mismatches 143; Indels 105; Gaps 12;
QY 3 ELHQP-----GNCDCLNGTGVSNKYFSNIHWCNPKKFGQHCIEIDSKTCYEGNGHFY 58
Db 74 QCHSVPRSCSEPRFCNGTQQQALYSDP--VCQPDGFGKRCIDTRATCFEGQGIT 132
QY 59 RGKASTDTMGPRCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPCYVQV 117
Db 133 RGTWTAENGAEICINWNSALSQKPSYARRPNAIKLGLGNHNYCRNPNDRDVKPCYVFXA 192
QY 118 GL-----KPLVQECMV----- 128

Db 193 GKTTTFCSTPACPKGFTEDCYVGKVTYRGTHSFTTSKASCLPWNMILGKTYTAWRA 252
QY 129 -----HDCADKXKFKQ-----CQCKTLR--PRFKIIG 154
Db 253 NSQALGLGRHNYCRNPDGDAKPMCHVMKDKLTWCEYCDMSPCSTCGLQVQKQPQFRIKGG 312
QY 155 EFTTIENQWFAAIY--RRHRGGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVVLGR 213
Db 313 LFTDIISHPMQAAIFVKNKRSPOGERFLCGGVLISSCWILSAAHCFVERPPPHLKVVLGR 372
QY 214 SRLNSNTQGMKEPEVENLILHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIOTICLP 273
Db 373 YRVVPGEEOKEFEKIVHKEFDODT--YNDIALLOKSDSRCAQESSWRTVCLPP 433
QY 274 MYNDPQF-----CTSCITGKENSVDLYPEOLKMTVVKLISHRECCOPHYGSEVTT 329
Db 431 ----DPDVLQPDWTCELSYGKHEASPPFFSRLKEAHVRLYPSRCSQHLNFKNTITS 486
QY 330 KMLCAADP-----QWKTDSQCDSGGPLVCSLQGRMTLTGIVSWRGCAKDKPGVYTRV 384
Db 487 NMLCAGDTRTGGNQDVHDACQDSGGPLVCMIDKMTLLGIISWGLGCGQKDPGVGITYKV 546
QY 385 SHELPMIRSHTK 397
Db 547 TNYLNMIQDNMKQ 559
RESULT 13
JS0597
t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0597
R:Kaezschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; D
Gene 185, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat D
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0597
A:Molecule type: mRNA
A:Residues: 1-477 <KRA>
A:Cross-references: GB:M63987; NID:G166070; PIDN:AAA1591.1; PID:G166071
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat ho
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <1FA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: kringle homology <KRG>
F:226-471/Domain: trypsin homology <TRY>
F:42-72, 70-79, 87-98, 92-109, 111-120, 128-209, 149-191, 180-204, 214-345, 257-273, 265-334, 359
F:153, 398/Binding site: carboxylate (Asn) (covalent) #status predicted
F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
F:272, 321, 428/Active site: His, Asp, Ser #status predicted
Query Match 37.2%; Score 840.5; DB 2; Length 477;
Best Local Similarity 43.0%; Pred. No. 1.2e-57;
Matches 175; Conservative 56; Mismatches 147; Indels 29; Gaps 10;
QY 5 HOVPSN-CD---CLNGTGVSNKYFSNIHWCNPKKFGQHCIEIDSKTCYEGNGHFYRG 60
Db 80 HTVPVNSCSEPRFCNGTQQAQVYSDP--VCQPDGFGKRCIEVTRATCYEGQVYRG 138
QY 61 KASTDTMGPRCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPCYVQVGLK 120
Db 139 TWSTAESRVEICINWNSLSLRTYNGRMPDAFNGLGNHNYCRNPNAPKAPWCYVIKAGK 198
QY 121 PLVQECMVHDCADGKLKFCQG--OKTLRPRFKIIGSEFTTIENQWFAAIYRHRGGS-VT 178
Db 199 FTSSECSVPVCS-----KATCGLRKYPQJHSTGLGTFDITSHPMQAAIFARNRSGSR 254
QY 179 YVCGSLISPCWISATHCFIDYPKKEDYI-----VYLGRSLNSNTQGMKEPEVENLIL 233

```

Db      255  FLCCGILISSCWLTAAHCF-----QESYLPDLQKVLGRYRVKPGSEEQTFVKXIV 309
QY      234  HKDYSADTLAHNDIALKIRSEGRCAQPSRTIOTICLPSMYNDPQGTSCBTGPOKE 293
Db      310  HKFEDDDT--YNDIALQLKSDSPQCAQESDSVRAICLPEANLQLPDWTCELSGYGKH 367
QY      294  NSTDYLYPEQKQMTVKLISHRECQCPHYGSEVTTKMLCAADPOWKT-----DSCQGD 347
Db      368  KSSSPFYSEQLKEGHVRLPSSRCAPKFLFNKVTNNWLCAGDRSGEIIYPNVHDACGD 427
QY      348  SGGPLVCSLQGRMTLTGIVSGRGALCKDKGQVYTRVSHFPLWRSH 394
Db      428  SGGPLVCMNDNRMTLLGIISGVGCGEXDVPQVTKVTNYLGWIRDN 474

RESULT 14
A29941
t-plasminogen activator (EC 3.4.21.68) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A29941; S48205; S48207; S48206
R:Rickles, R.J.; Darrow, A.L.; Strickland, S.
J. Biol. Chem. 263, 1563-1569, 1988
A:Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator m
A:Reference number: A29941; MUID:86087303; PMID:2826484
A:Accession: A29941
A:Molecule type: mRNA
A:Residues: 1-559 <RIC>
A:Cross-references: GB:J03520; NID:G202109; PIDN:AAA40470.1; PID:G202110
R:Liinen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A:Title: Characterization of the murine plasma fibrinolytic system.
A:Reference number: S48202; MUID:95010076; PMID:7523120
A:Accession: S48205
A:Molecule type: protein
A:Residues: 33-37,'X',39-40 <LIJ>
A:Accession: S48207
A:Molecule type: protein
A:Residues: 309-316 <LI2>
A:Accession: S48206
A:Molecule type: protein
A:Residues: 33-37,'X',39-40 <LIW>
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-29/Domain: propeptide #status predicted <PRO>
F:30-559/Product: t-plasminogen activator #status predicted <MAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F:38-75/Domain: fibronectin type I repeat homology <IFI>
F:83-116/Domain: EGF homology <EGF>
F:124-205/Domain: kringle homology <KRI>
F:213-294/Domain: kringle homology <KRI2>
F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F:309-553/Domain: trypsin homology <TRY>
F:38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4
F:149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-116 (plasmin, trypsin) #status predicted
F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 36.8%; Score 829.5; DB 1; Length 559;
Best Local Similarity 35.7%; Pred. No. 9.9e-57;
Matches 176; Conservative 63; Mismatches 149; Indels 105; Gaps 9;

QY      3  ELHQP-----SNCDLNGGTCVSNKFSNIHWNCNPKKFGQCEIDKSKTCYEGNHFY 58
Db      74  QCHSVPRSSSEPRCFNGTCCQALYSDP--VCQCPDGFVKGKCDIDTRATCFEEQITY 132
QY      59  RGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRRPWCYVQVG 118
Db      133  RGTWSTAESAECINWNSSVLSLKPYNARRPNAIKLGLGNHNYCRNPDRLDKPWCYVFXA 192
QY      119  LKPLVQECMVHDCADGLK----- 137

```

```

Db      193  GKTTTEFCSTPACPKGSEDCYVGKGYRGTGTHSLTTSQASCLPWNISVLMGKSYTAWRT 252
QY      138  -----FQCG-QKTLRPRFKIIG 154
Db      253  NSQALGLARHNYCRNPDGARPWCHVMKORKLWEYCDMSPCSTCGLRQYKRPOFRKGG 312
QY      155  EFTTIENQPFALY--RRHRGSGVTVYCGSLISPCWVISATHCFIDYPKKEDIYVVLGR 213
Db      313  LYTDITSHPQAAIFVYNKRSPGERFLCGGVLTSSCWLSAAHCFLEFFPNHLKVVVLR 372
QY      214  SRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSEGRCAQPSRTIOTICLP 273
Db      373  TYRVVPGEEBQTEIEKIYVHEEFDODT--YNDIALQLRQSOKCAQESSVGTACL 430
QY      274  SMYNDPQF-----GTSCEITGFGKENSTDYLPQLKMTVVKLISHRECQCPHYGSEVTT 329
Db      431  ----DNLQLPDWTECELSGYGKHEASSPFFSDRLKEAHVRLYPSRCTSHLFNKVTN 486
QY      330  KMLCAADP-----QWKTDSQCGSGPLVCSLQGRMTLTGIVSGRGCALCKDKPGVYTRV 384
Db      487  NMLCAGDTRSGGNQDLHDACQSGSGPLVCMINKQMTLTGIIISWGLCGGCKDVPGVYTKV 546
QY      385  SHFLPWIRSHKE 397
Db      547  TNYLDWIHDNMKQ 559

RESULT 15
JS0600
t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0600
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; D.
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat D.
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0600
A:Molecule type: mRNA
A:Residues: 1-394 <KRA>
A:Cross-references: GB:M63990; NID:G16078; PIDN:AAA1595.1; PID:G166079
A:Note: the authors translated the codon ATC for residue 75 as Thr
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-334/Product: plasminogen activator gamma #status predicted <PLA>
F:45-126/Domain: kringle homology <KRG>
F:43-388/Domain: trypsin homology <TRY>
F:45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds
F:142-143/Cleavage site: His-Ser (plasmin) #status predicted
F:189,238,345/Active site: His, Asp, Ser #status predicted
F:315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 33.5%; Score 756; DB 2; Length 394;
Best Local Similarity 42.9%; Pred. No. 3.4e-51;
Matches 155; Conservative 56; Mismatches 134; Indels 16; Gaps 7;

QY      45  DKSKTCTYEGNHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRN 104
Db      40  DPHATCYKQGVYRGTWSTSSGACQINWNLSLRITNGRMPEAVKLGHNHNYCRN 99
QY      105  PNRRRPWCYVQVGLKPLVQECMVHDCADGLKFCQG-QKTLRPRFKIIGSEFTTIENOP 163
Db      100  PDGASKPWCYVIKARKFTSESQSVPCS---KATCGLRKYKEPQLHSTGGFLTDTISHP 155
QY      164  WFAALYRRHRGGS-VTVYCGGSLISPCWVISATHCFID-YPKKEDIYVYLGSRNSNTQ 221
Db      156  WQAAIFAQNRSSGRRFLCGGILISSCWLTAAHCFQERYPPQHLRVV-LGRTYRVKPGK 214
QY      222  GEMKFEVENLILHKDYSADTLAHNDIALLKIRSEGRCAQPSRTIOTICLPMSMYNDPQF 281

```

```
Db 215 EOTFEVEXICVHEEFDDDT--YNDIALQLKSGSPQCAQESDSVRAICLPEANLQLPD 272
QY 282 GTCEITGKGENSTDYLYPEQLKMTWKILSHRECOQPHYGVSEVYTKMLCAADPOWKT 341
Db 273 WTECELSGYGKXSSPFFSEQLKEGHVLYPSSRCTSKFLPNKVTNNMLCAGTRSGE 332
QY 342 -----DSQGDGGGLVCSLQGRMTLTGIVSWGRGCAKDKPGVVYTRVSHFLPWSHT 395
Db 333 IYVNHVHACQGDGGGLVCMNDNHNMTLLGIISMGVCGGEKDPGVYVTKVTNYLGMIRDNM 392
QY 396 K 396
Db 393 R 393

RESULT 16
A46688
hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human
C:Species: Homo sapiens (man)
C>Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
C:Accession: A46688
R: Miyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
J. Biol. Chem. 268, 10024-10028, 1993
A: Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease
d coagulation factor XII.
A: Reference number: A46688, MUID: 93252878; PMID: 7683665
A: Accession: A46688
A: Molecule type: mRNA
A: Residues: 1-655 <MIY>
A: Cross-references: DDBJ: D14012; NID: G219680; PIDN: BAA03113.1; PID: G219681
A: Experimental source: liver (mRNA); serum (protein)
A: Note: sequence extracted from NCBI backbone (NCBIN: 131227, NCBIPI: 131228)
C: Geneticks:
A: Gene: GDB: HGFA; HGFA; HGFAP
A: Cross-references: GDB: 9954514
A: Map position: 4p16-4p16
C: Function:
A: Description: activates hepatocyte growth factor by specific proteolytic cleavage
A: Pathway: tissue repair and regeneration
C: Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C: Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase
F: 1-34/Domain: signal sequence #status predicted <SIG>
F: 108-148/Domain: fibronectin type II repeat homology <IF2>
F: 164-197/Domain: EGF homology <EG1>
F: 202-237/Domain: fibronectin type I repeat homology <IF1>
F: 245-278/Domain: EGF homology <EG2>
F: 286-367/Domain: kringle homology <KRG>
F: 373-407/Product: hepatocyte growth factor activator light chain #status experimental <
F: 408-655/Product: hepatocyte growth factor activator heavy chain #status experimental <
F: 408-641/Domain: trypsin homology <TRY>
F: 40, 48, 290, 468, 492, 546/Binding site: carbohydrate (Asn) (covalent) #status predicted
F: 164-175, 169-186, 188-197, 202-230, 228-237, 245-256, 250-267, 269-278, 286-367, 307-349, 338-36
F: 447, 497, 598/Active site: His, Asp, Ser #status predicted

Query Match 32.2%; Score 726.5; DB 1; Length 655;
Best Local Similarity 37.2%; Pred. No. 1.2e-48;
Matches 158; Conservative 53; Mismatches 161; Indels 53; Gaps 9;
QY 5 HQVPSNCLNGGTGCVSNKYFSNIHW-----CNCPKKGQGHCEIDKSKTCYEGNGH 56
Db 242 HTACUSSCLNGGTG-----HLIVATGTTVCACPPGFAGRLCNTEPDERCFLNGT 292
QY 57 FYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQLGLGHNYCRNPDRNRPPWCYV- 115
Db 293 GYRGVASTSASGLSLCLAWNSDLLYQELHVDVSGAALLGLGPHAYCRNPDRNDPWCYV 352
QY 116 -----QVGLKPLVQECMVHDCADGKLFQCGQK-----TLRPFKI 151
Db 353 KDSALSWEYCRLEACESITRVQLSPDLLATLPEPASGRQ--ACGRHRKXKTLFLRPR--I 408
QY 152 IGGTEFTIENQFWFAAIIYRRHRGSGVTVYVCGSLISPCWVISATHCFIDYPKKEDIYVL 211
```

```
Db 409 IGGSSSLPGSHPLAAIY---IGDS---FCAGSLVHTCWVVSAAHCFESHSPRSVSVL 462
QY 212 GRSRLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSKRGCAQPSFTIQTIC 271
Db 463 GQHFNFRTDVTQTFGIEKIPYTLTVSVFNPDSH-DLVLIIRLKKKKGRCATRSQVQPIC 521
QY 272 LPSMYNDPQFGTSCIEITGFGKENSTDYLYPEQLKMTWKILSHRECOOPHYGVSEVYTKM 331
Db 522 LPEGSFTFPAGHKCQIAGWGLHDENVSGYSSSLREALVPLVADHKCSPEVYGADISPNM 581
QY 332 LCAADPOWKTDSCQDGGGLVCSLQGRMTLTGIVSWGRGCAKDKPGVVYTRVSHFLPWI 391
Db 582 LCAGYFDCKSDACQDGGGLVCKNGVAYLYGIISWGDGCGRLHKPGVYTRVANYVDWI 641
QY 392 RSHTK 396
Db 642 NDRIR 546

RESULT 17
S28941
coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)
N: Alternate names: Hageman factor
C: Species: Cavia porcellus (guinea pig)
C: Date: 25-Feb-1994 #sequence_revision 03-Aug-1995 #text_change 21-Jan-2000
C: Accession: S28941
R: Senba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambara, T.; Okabe, J.
Biochim. Biophys. Acta 1159, 113-121, 1992
A: Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage
A: Reference number: S28941, MUID: 93003367; PMID: 1390917
A: Accession: S28941
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-603 <SEM>
A: Cross-references: EMBL: X68615; NID: G49578; PIDN: CAA48600.1; PID: G49579
C: Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology
C: Keywords: hydrolase; serine proteinase
F: 46-87/Domain: fibronectin type II repeat homology <IF2>
F: 134-169/Domain: fibronectin type I repeat homology <FBI>
F: 177-208/Domain: EGF homology <EGF>
F: 216-294/Domain: kringle homology <KRG>
F: 359-597/Domain: trypsin homology <TRY>

Query Match 31.0%; Score 700.5; DB 2; Length 603;
Best Local Similarity 36.7%; Pred. No. 1.1e-46;
Matches 160; Conservative 60; Mismatches 147; Indels 69; Gaps 13;
QY 13 CLNGGTCVSNKYFSNIHWNCNPKKFGQGHCEIDKSKTCYEGNGHFGYRGKASTDTMGRPCL 72
Db 182 CLNGGRGLE---VEGHHLCDCPMGYTGPFCDLDTTASCYEGRGVSIRGMARTTVSGAKCQ 238
QY 73 PWSATVLQOQTYHAHRS-DALQGLGHNYCRNPDRNRPPWCYVQVGLKPLVQECMVHDC 131
Db 239 RWAS-----EATYRNMTAEQALRRGLGHTFCRNPDPNDTRPWCYVWGNRLSWEYCDLAQC 294
QY 132 -----ADGKLKFK-----CGKTLRPRF 149
Db 295 QYPPQPTATPHDFEPKXPSSSLSLQTPQPTQNALANELPETSLLCGQR-LRKRL 353
QY 150 -----KIIGGEFTTIENQFWFAAIIYRRHRGSGVTVYVCGSLISPCWVISATHCFIDYPKKE 205
Db 354 SLSLRVGGVLVALPGAHPYIALLY-----NFSGSLIAPCWLTAAHCLQNPAPR 407
QY 206 DYIVYIGRRLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSKRGCAQPS 264
Db 408 ELKVLQGDQRHNSQCHCQTLAVHSYRLHEAFSPS--SYLNDLALLRLQKSDGSAQLS 465
QY 265 RTIQTICLPMYNDPQFG--TSCIEITGFGKENSTDYLYPEQLKMTWKILSHRECOOPHY 322
Db 466 PYQTVCLPSGAPPSSEITTCCEVAGWGHQFEGABEYSSFLQEAQVPLISSERCSSEV 525
QY 323 YGSEVTKMLCAADPOWKTDSCQDGGGLVCKNGVAYLYGIISWGDGCGRLHKPGVYTRVANYVDWI 379
```

```

Db      526 HGDAFUSGMLCAGFEGGTDACQSGGFLVCEDEAAEHRLTLRGIVSWGSGGRNPKG 585
QY      380 VYTRVSHFLPWIRSH 395
      586 VYTDVASYLTWQKHT 601

RESULT 18
KPFU12
coagulation factor XIIa (EC 3.4.21.38) precursor [validated] - human
N:Alternate names: Hageman factor (activated)
C:Species: Homo sapiens (man)
C:Date: 27-Nov-1985 #sequence_revision 30-Jun-1991 #text_change 08-Dec-2000
C:Accession: A29411; A26814; A00930; A25191; A22448; A21037
R:Coil: D.E.; MacGillivray, R.T.A.
J. Biol. Chem. 262, 13662-13673, 1987
A:Title: Characterization of the human blood coagulation factor XII gene. Intron/exon ge
A:Reference number: A29411; MUID:88007593; PMID:2888762
A:Accession: A29411
A:Molecule type: DNA
A:Residues: 1-615 <CO>
A:Cross-references: GB:M17466; GB:J02807; NID:gl80355; PIDN:AA59490.1; PID:gl80357
R:Tripos, M.; Citarella, F.; Guida, S.; Galeffi, P.; Fantoni, A.; Cortese, R.
Nucleic Acids Res. 14, 3146, 1986
A:Title: cDNA sequence coding for human coagulation factor XII (Hageman).
A:Reference number: A26814; MUID:86176794; PMID:3754331
A:Accession: A26814
A:Molecule type: mRNA
A:Residues: 4-615 <TRI>
A:Cross-references: GB:M31315; NID:gl82291; PIDN:AAA70225.1; PID:gl82292
R:Coil, D.E.; Edgell, C.J.S.; Louie, G.V.; Zoller, M.J.; Brayer, G.D.; Macgillivray, R.T.
J. Biol. Chem. 260, 13666-13676, 1985
A:Title: Characterization of human blood coagulation factor XII cDNA. Prediction of the
A:Reference number: A00930; MUID:86033830; PMID:3877053
A:Accession: A00930
A:Molecule type: mRNA
A:Residues: 14-332, 'S', 334-615 <CO2>
A:Cross-references: GB:M11723; NID:gl80358; PIDN:AAA51986.1; PID:gl80359
R:Que, B.G.; Davie, E.W.
Biochemistry 25, 1525-1528, 1986
A:Title: Characterization of a cDNA coding for human factor XII (Hageman factor).
A:Reference number: A25191; MUID:86216049; PMID:3011063
A:Accession: A25191
A:Molecule type: mRNA
A:Residues: 146-378, 'G', 380-615 <QUE>
A:Cross-references: GB:M31314; NID:gl80360; PIDN:AAA70224.1; PID:gl80361
R:McMullen, B.A.; Fujikawa, K.
J. Biol. Chem. 260, 5328-5341, 1985
A:Title: Amino acid sequence of the heavy chain of human alpha-factor XIIa (activated Ha
A:Reference number: A22248; MUID:85182674; PMID:3886654
A:Accession: A22248
A:Molecule type: protein
A:Residues: 20-379 <MCM>
R:Fujikawa, K.; McMullen, B.A.
J. Biol. Chem. 258, 10924-10933, 1983
A:Title: Amino acid sequence of human beta-factor XIIa.
A:Reference number: A21037; MUID:83291041; PMID:6604055
A:Accession: A21037
A:Molecule type: protein
A:Residues: 354-362:373-615 <FUJ>
R:Harris, R.J.; Ling, V.T.; Spellman, M.W.
J. Biol. Chem. 267, 5102-5107, 1992
A:Title: O-linked fucose is present in the first epidermal growth factor domain of facto
A:Reference number: A44606; MUID:92184750; PMID:1544894
C:Genetics:
A:Contents: annotation; carbohydrate binding site
A:Gene: GDB:F12
A:Cross-references: GDB:119892; OMTM:234000
A:Map position: Sq34-5qter
A:Introns: 19/3; 39/1; 72/2; 96/1; 133/1; 177/1; 212/1; 267/2; 340/1; 417/2; 463/1; 511/
C:Complex: factor XII, prekallikrein, and HMW kininogen form a complex bound to anionic
C:Function:

```

```

A:Description: factor XIIa catalyzes the proteolytic activation of plasminogen, plasma
ikrein
A:Pathway: blood coagulation; fibrinolysis
C:Superfamily: coagulation factor XII, EGF homology; fibronectin type I repeat homolog
C:Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolyase; kringle; plasma;
F1-19/Domains: signal sequence #status predicted <SIG>
F120-372,373-615/Product: coagulation factor XIIa, alpha form #status experimental <A1
F147-88/Domains: fibronectin type II repeat homology <FB2>
F198-130/Domains: EGF homology <EG1>
F1135-170/Domains: fibronectin type I repeat homology <IF1>
F1178-209/Domains: EGF homology <EG2>
F1217-295/Domains: kringle homology <KR>
F1298-356/Region: proline-rich
F1354-362,373-615/Product: coagulation factor XIIa, beta form #status experimental <B1
F1373-609/Domains: trypsin homology <TRY>
F198-110,104-119,121-130,135-163,161-170,178-189,183-198,200-209,217-295,238-277,266-2
F1249,433/Binding site: carbohydrate, (Thr) (covalent) #status experimental
F1299,305,328,329,337/Binding site: carbohydrate (Asn) (covalent) #status experimental
F1308/Binding site: carbohydrate (Ser) (covalent) #status predicted
F1412,461,563/Active site: His, Asp, Ser #status predicted

Query Match      30.24;      Score 681;      DB 1;      Length 615;
Best Local Similarity 34.18;      Pred No. 3.7e-45;
Matches 152;      Conservative 58;      Mismatches 158;      Indels 78;      Gaps 10;

QY      13 CLNGGTCVSNKYFSNIHMCNCPKFGQGHCEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
      183 CLHGGRCLE---VEGHLCHCPVGYTGFCFDDVDTKASYCDGRGLSYRGLARTTLTGAPCQ 239
QY      73 PMSATVLOQTY-HAHRSDALQGLGHXNYCRPNDRRRPWCYVQGLKPLVQECMVHDC 131
      240 PWAS-----EATYRNVTAEQARNWGLGHAFCRPNDRPNDIRPWCFLNDRILSWEYCDLAQC 295
QY      132 -----ADGKL----- 136
      296 QTPTQAAPPTVSPRLHVLMPAQAPPKQPQPTTRTPQSGTQALPAKREQPSLTRNG 355
QY      137 KFQCGQ---KTLRPRFKLIIGETTTIENQPFVAIYRRHGGSVTVYVCGGSLISPCWVIS 193
      356 PLSCGQRLKSLSSMTVRVVGGLVALRGAHPYIAALYMGHS-----FCAGSLIAPCWVLT 409
QY      194 ATHCFIDYPKEDYIVVGLSRSLNSNTQGMKEFEVENLILKDYSDATLAHNDIALLLKI 253
      410 AAHLQDRPAPEDLTVLVGQERRNHSCPEQTLAVRSYRLHEAFS--PVSYQHDALLRL 467
QY      254 R-SKEGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGKENSTDYLYPEQLKMTVVVKLI 312
      468 QEDADGSCALLSPVQPVCLPSGAARPESETLLCQVAGWGHPGAGAEYASFLQEAQVPFL 527
QY      313 SHRECQPHYGVSEVTKMLCAADPOWKTDSCGDSGGPLVCSLQG---RMTLTGIVSWG 369
      528 SLERCAPDVHGSIIPLGMLCAGFLEGGTDACQSGGGLVCEQDAARRLTLQGIISWG 587
QY      370 RGCALDKDPGVYTRVSHFLPWIRSH 395
      588 SGCGDRNKPQVYTDVAYYLAWEHT 613

RESULT 19
JC5878
plasma hyaluronan-binding protein precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 16-Jul-1999
C:Accession: JC5878
R:Hashimoto, K.; Tobe, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura,
Biol. Pharm. Bull. 20, 1127-1130, 1997
A:Title: Cloning of the cDNA for a mouse homologue of human PHBP: A novel hyaluronan-b
A:Reference number: JC5878; MUID:9805239; PMID:9401717
A:Accession: JC5878
A:Molecule type: mRNA
A:Residues: 1-558 <HAS>
C:Comment: This protein acts as serine protease.

```

C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsin
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-311/Product: plasma hyaluronan-binding protein large chain #status predicted <NATL>
F:75-106/Domain: EGF homology <EG1>
F:113-145/Domain: EGF homology <EG2>
F:152-185/Domain: EGF homology <EG3>
F:192-274/Domain: kringle homology <KRI>
F:312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <NATL>
F:312-548/Domain: trypsin homology <TRY>

Query Match 29.4%; Score 664.5; DB 2; Length 558;
Best Local Similarity 36.9%; Pred. No. 6.4e-44;
Matches 154; Conservative 57; Mismatches 163; Indels 43; Gaps 14;
Qy 13 CLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHYRKGKASTDTWGRPCL 72
Db 157 CONGGVCSRRHRSRFTCACPQYKGFCEIGFD-DCVVGDSYRGKVKTVNQNPCL 214
Qy 73 PMSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQGLKPL-----VQEC 126
Db 215 YNSHLLLOQTYNMFEDAEATHGAEHNFRCNPDGDKPWCYKVNSEKVKWEYCDVTV 274
Qy 127 MVHDCADGKLF-----QCQKTLRPRF-KIIGSEFTTIENQWFAAIYRHR 169
Db 275 PVDPNPNVSLLEPVMELPFESCGKTEVAEHAHVRIYGGFKSTAGKHPWQVSLQSLP 334
Qy 170 ---RRHGGSVTVYVCGSLISPCWISATHCFIDYPKKEDIYVYLGSRSLNSNTQEMKF 226
Db 335 LTTSMPQG---HFCGGALHPWCWLTAHC-TDINTKHLKV-LGDQDLKKTSHQTF 388
Qy 227 EVENLIHKDYADTLAHNDIALKIRSKGRCAPQSRITQICLPSMYNDPQF--GT 283
Db 389 RVEKILKYSQVNERDEIPNDIALKLPVGHCALESRYKVTVCPLPS---DFFPSGTEC 445
Qy 286 EITGFGKENSVDYLPQKMTVVKLISHRECQCPHYGVSEVTTKMLCAADPQWK-TDSC 344
Db 446 HISGVTETGE--GSRQLLDKAVKLIANPLCNSRQLYDHTIDDSMICAGNLQKPGSDTC 503
Qy 345 QGDSGGPLVCSLQGRMTLTGIVSWGRGCALKDEGVYTRVSHFLPWIRSHTKKEENG 401
Db 504 QGDSGGPLTCEKDGTYVYGVISWGLECG--KRGVYTVQTKPLNWKATIKESG 559

RESULT 20
JC4735
Plasma hyaluronan-binding protein precursor - human
N:Alternate names: hepatocyte growth factor activator-like protein; PHBP
N:Contains: serine proteinase (EC 3.4.21.-)
C:Species: Homo sapiens (man)
C>Date: 15-Oct-1995 #sequence_revision 16-Aug-1996 #text_change 19-Jul-2002
C:Accession: J04795
R:Choi-Miura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Mazda, T.; Tomita, M.
J. Biochem. 119, 1157-1165, 1996
A:Title: Purification and characterization of a novel hyaluronan-binding protein (PHBP)
r activator.
A:Reference number: J04795; MUID:96425001; PMID:8827452
A:Accession: J04795
A:Molecule type: mRNA
A:Residues: 1-560 <CHO>
A:Cross-references: GB:S83182; NID:G1836158; PIDN:AAB46909.1; PID:G1836159
A:Experimental source: plasma
A:Note: parts of this sequence, including the amino ends of the mature chains, were deter
C:Genetics:
A:Gene: GDB:HABP2; HABP; PHBP; HGFBAL
A:Cross-references: GDB:4573962
C:Complex: a disulfide-bonded heterodimer of chains produced from the same precursor; th
C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsin
C:Keywords: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsin
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-311/Product: plasma hyaluronan-binding protein, 50K chain #status predicted <50K>
F:77-106/Domain: EGF homology <EG1>
F:113-147/Domain: EGF homology <EG2>
F:154-187/Domain: EGF homology <EG3>

F:194-276/Domain: kringle homology <KRI>
F:314-550/Domain: trypsin homology <TRY>
F:314-516/Product: plasma hyaluronan-binding protein, catalytic chain #status predicted
F:54-207/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:77-88, 82-97, 99-108, 115-125, 120-136, 138-147, 154-165, 159-176, 178-187, 194-276, 215-257, 2
F:362, 405, 509/Active site: His, Asp, Ser #status predicted

Query Match 28.9%; Score 651.5; DB 1; Length 560;
Best Local Similarity 36.4%; Pred. No. 6.6e-43;
Matches 152; Conservative 58; Mismatches 161; Indels 47; Gaps 14;
Qy 13 CLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHYRKGKASTDTWGRPCL 72
Db 159 CONGATCSRHRKRSKFTCACPDQFKGFCEIG-SDDCVVGDSYRGKVNRTVQHAACL 216
Qy 73 PMSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQGLKPLVQEC-CMVHDC 131
Db 217 YNSHLLLOQTYNMFEDAEATHGAEHNFRCNPDGDKPWCYKVNSEKVKWEYCDVTV 276
Qy 132 ADGKLF-----QCQKTLRPRF-KIIGSEFTTIENQWFAAIYRHR 173
Db 277 SAQDVAPBESPTPESTKLPFGDSCGKTEIAERKIKRIYGVGFKSTAGKHPWQASLQ---- 332
Qy 174 GGSVT-----YVCGSLISPCWISATHCFIDYPKKEDIYVYLGSRSLNSNTQEMK 225
Db 333 -SSLPTISMPQGHFCGGALHPWCWLTAHC-TDI-KTRHLKVLGDODLKKKEEFHES 389
Qy 226 FEVENLIHKDYADTLAHNDIALKIRSKGRCAPQSRITQICLPSMYNDPQF--GT 283
Db 390 FRVEKIFKYSHYNERDEIPNDIALKLPVGHCALESRYKVTVCPLPS---DGSFPGS 445
Qy 284 SCITGFGKENSVDYLPQKMTVVKLISHRECQCPHYGVSEVTTKMLCAADPQWK-TD 342
Db 446 ECHISGVTETGK--GSRQLLDKAVKLIANPLCNSRQLYDHTIDDSMICAGNLQKPGD 503
Qy 343 SCQDSSGGLVCSLQGRMTLTGIVSWGRGCALKDEGVYTRVSHFLPWIRSHTKKEENG 400
Db 504 TCQDSSGGLTCEKDGTYVYGVISWGLECG--KRGVYTVQTKPLNWKATIKESG 559

RESULT 21
S45281
Coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)
N:Alternate names: Hageman factor (activated)
C:Species: Bos primigenius taurus (cattle)
C>Date: 10-Apr-1995 #sequence_revision 22-Apr-1995 #text_change 21-Jan-2000
C:Accession: S45281; A61329
R:Shibuya, Y.; Semba, U.; Okabe, H.; Kanbara, T.; Yamamoto, T.
Biochim. Biophys. Acta 1206, 63-70, 1994
A:Title: Primary structure of bovine Hageman factor (blood coagulation factor XII): co
A:Reference number: S45281; MUID:94242782; PMID:8186251
A:Accession: S45281
A:Molecule type: mRNA
A:Residues: 1-593 <SHI>
A:Cross-references: GB:S70164
A:Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue 70
is Pro, CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue 286 as
is, and ATC for residue 505 as Leu
R:Fujikawa, K.; Walsh, K.A.; Davie, E.W.
Biochemistry 16, 2270-2278, 1977
A:Title: Isolation and characterization of bovine factor XII (Hageman factor).
A:Reference number: A61329; MUID:77182112; PMID:861210
A:Accession: A61329
A:Molecule type: protein
A:Residues: 10-16, 'X', 18-19, 525-550 <FUJ>
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homolog
C:Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma;
F:37-78/Domain: fibronectin type II repeat homology <IF2>
F:88-120/Domain: EGF homology <EGF>
F:125-160/Domain: fibronectin type I repeat homology <FBI>
F:207-287/Domain: kringle homology <KRG>
F:350-587/Domain: trypsin homology <TRY>
F:541/Active site: Ser #status predicted

Query Match 28.3%; Score 638; DB 2; Length 593;
 Best Local Similarity 34.2%; Pred. No. 7.8e-42;
 Matches 151; Conservative 59; Mismatches 164; Indels 68; Gaps 13;

QY 6 QVPSNCDCLNGTGVSNKYFNIHWCNPKKFGQCHCEIDSKTCYE--GNGHFYRGKAS 63
 DB 166 QVCRNTPCLNGSDCLQAE--GHRLCRCAPSFAGRLCDVDLKASCYDDRDRLGSLYRGWAG 222

QY 64 TDTMRPCLPWSATVLOQTY-HAHRSDALQGLGKXNYCENPDNRBRPCYVQVGLKPL 122
 DB 223 TLTSGAPCQSNAS---EATYVNTAEQVLNWLGLDHAFRCRNPNDTRPWFICWKGDRLS 278

QY 123 VOECWHDG--ADGKLKQF-----CGQ 142
 DB 279 WNYCRLAPCAAGHEHPPLSPSALQKPESTTQTLPSLTSGWCSPTPLASGPGPGCGQ 338

QY 143 ---KTLRPRKILIGEFTHIENQWFAIYRHRGGSVTYVCGSLISPCWVWISATHCFI 199
 DB 339 RLKRWLSLNRVGVGLVALPGAHPIYIAALYWDQ-----HFCAGSLIAPCWVLTAAHCLQ 392

QY 200 DYPKKEDYIVYLGSRSLNSNTQGMKFEVENLIHKDYSADTLAHNDIALLKIR-SKEG 258
 DB 393 NRPAKELTVLQGDRHNSCEQCOTLAVRDYRLHEAFSPITYQH--DLALVRLQESADG 450

QY 259 RCAPSRITQICLPSMYNDPQFTS--CEITGFKENSTDYLPQOLKMTVVKLISHRE 316
 DB 451 CCAHPSFPVQVCLPSTAAARPAESEAACEVAGWHQFEGGE-YSSFLQEAQVPLIDPQR 509

QY 317 CQOPHYTGSEVTTKMLCAADPQWKTDCSQGDSGGPLVC---SLQGRMTLGIYVSWGRGCA 373
 DB 510 CSAFDVHGAARTQMLCAGFLEGGTDACQDGGPLVCEDETPERQILGIYVSWGGCG 569

QY 374 LKDRPGYTVRSHPLMIRSHT 395
 DB 570 NLRKPGYTVDVANYLAWIREHT 591

RESULT 22

PLPG
 Plasmin (EC 3.4.21.7) precursor - pig (fragment)
 N:Alternate names: plasminogen
 N:Contains: miniplasminogen
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 07-Sep-1990 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
 C:Accession: S03733; S03737; A25834
 R:Schaller, J.; Marti, T.; Roesele, S.J.; Kaempfer, U.; Rickli, E.E.
 A:Title: Amino acid sequence of the heavy chain of porcine plasmin. Comparison of the ca
 A:Reference number: S03733
 A:Accession: S03733
 A:Molecule type: protein
 A:Residues: 1-560 <SCH>
 R:Brundisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;
 Eur. J. Biochem. 114, 465-470, 1981
 A:Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
 A:Reference number: S03735; MUID:81212097; PMID:7238497
 A:Accession: S03737
 A:Molecule type: protein
 A:Residues: 1-57 <BRU>
 R:Marti, T.; Schaller, J.; Rickli, E.E.
 Eur. J. Biochem. 149, 279-285, 1985
 A:Title: Determination of the complete amino-acid sequence of porcine miniplasminogen.
 A:Reference number: A25834; MUID:85203907; PMID:3846533
 A:Accession: A25834
 A:Molecule type: protein
 A:Residues: 450-790 <MAR>
 C:Function:
 A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va
 ns the walls of the graafian follicle; also activates the urokinase-type plasminogen act
 C:Superfamily: fibrinolysis
 C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasmin; plasma; serine prote

F:1-790/Product: plasminogen #status predicted <PRO>
 F:1-77/Domain: plasminogen-related protein precursor homology (fragment) <PLPH>
 F:1-77/Domain: activation peptide #status predicted <APT>
 F:78-560/Product: plasmin chain A #status predicted <ACH>
 F:84-162/Domain: kringle homology <KR1>
 F:166-243/Domain: kringle homology <KR2>
 F:256-333/Domain: kringle homology <KR3>
 F:358-435/Domain: kringle homology <KR4>
 F:450-790/Product: miniplasminogen #status experimental <MIN>
 F:461-540/Domain: kringle homology <KR5>
 F:561-790/Product: plasmin chain B #status experimental <BCH>
 F:561-783/Domain: trypsin homology <TRY>
 F:30-54,34-42,84-162,105-145,133-157,166-243,169-297,187-226,215-238,256-333,277-316,31
 bonds: #status predicted
 F:602,645,740/Active site: His, Asp, Ser #status predicted

Query Match 22.5%; Score 508.5; DB 1; Length 790;
 Best Local Similarity 36.2%; Pred. No. 1.2e-31;
 Matches 130; Conservative 45; Mismatches 141; Indels 43; Gaps 14;

QY 45 DSKTCYEGNGHFYRGKASTDTMRPCLPWSATVLOQTYHAHR-----SDALQGLGKH 100
 DB 456 DLSEDCMFGNGKRYGRKRAITVAGVPCQEWAA-----QEPHRSIPTPTNPRAGLEK-N 509

QY 101 YCRNPD-NRRRPMCVYVQGLKPLVQECMVHDCADGKLFQCCQKTLRPR---FKIIGGEF 156
 DB 510 YCRNPDGDDGWCYT-TNPQKLFVDCVPCQVTS--SFDCKGPKVPEKKCPARVVGCV 566

QY 157 TTENCPFAAIYRHRGSGSVTYVCGSLISPCWVWISATHCFIDYPPKEDYIVYLGSRSL 216
 DB 567 SIPHSPWQISLURYRYRG---HFCGGTLISPEWLTAXHCKLEKSSSPSSYKVILGAHEE 622

QY 217 NSNTQGMKFEVENLIHKDYSADTLAHNDIALLKIRKEGRCQAPSRITQICLPSMY 276
 DB 623 YHLGEGVQIDVSKLF--KEPS-----EADIALKLSPP-----AVITKVIPLACLT-- 568

QY 277 NDPQP-----GTSCEITGFKENSTDYLPQOLKMTVVKLISHRECQPHYGVSEVTTKML 332
 DB 669 --PNYVADRTACYITGWGETGT--YGAGLLKEARLPVIENKVCNRYELGKVPSPNEL 724

QY 333 CAADPQWKTDCSQGDSGGPLVCSLQGRMTLGIYVSWGRGCAKDKPGYTVRSHPLFI 391
 DB 725 CAGHLAGGIDSCQDGGPLVCFKDKYLOQVTSWGLGCAFPNRPVTVRSRFTWI 783

RESULT 23

B61545
 Plasmin (EC 3.4.21.7) precursor - sheep (fragments)
 N:Alternate names: plasminogen
 N:Contains: miniplasminogen
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C>Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 17-Mar-1999
 C:Accession: B61545; S28200
 R:Schaller, J.; Rickli, E.E.
 Enzyme 40, 63-69, 1988
 A:Title: Structural aspects of the plasminogen of various species.
 A:Reference number: A61545; MUID:89005015; PMID:3168975
 A:Accession: B61545
 A:Molecule type: protein
 A:Residues: 1-37;38-117 <SCH>
 R:Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
 Protein Seq. Data Anal. 5, 21-25, 1992
 A:Title: Complete amino acid sequence of ovine miniplasminogen.
 A:Reference number: S28200; MUID:93149995; PMID:1492092
 A:Accession: S28200
 A:Molecule type: protein
 A:Residues: 118-460 <SC2>
 C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolo
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasmin; plasma; serine proteinase;
 F:1-37,38-117,118-460/Product: plasminogen (fragments) #status experimental <PRO>
 F:1-37/Domain: activation peptide (fragment) #status experimental <APT>
 F:38-117,118-230,231-460/Product: plasmin (fragments) #status experimental <MAT>
 F:41-118/Domain: kringle homology <KR4>

F:118-460/Product: miniplasminogen #status experimental <MIN>
F:132-211/Domains: kringle homology <KR>
F:226-460/Domains: plasmin chain B #status experimental <BC>
F:231-453/Domains: trypsin homology <TRY>
F:272,315,410/Active site: His, Asp, Ser #status predicted

```
Query Match      22.2%   Score 502;   DB 2;   Length 460;
Best Local Similarity 34.7%   Pred. No. 2,1e-31;
Matches 131;   Conservative 43;   Mismatches 158;   Indels 46;   Gaps 14;

QY  28  IHW--CNCPKKFGGHCIDSKTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYH 85
    :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  108 VRWEEFCNLKAPQAPSVENPEADCMLGIGKGYRGKATTVAGVPCQEWAA-----QEPH 162

QY  86  AH---RSDALQLGLGKKNYCRNPD--NRRRRPWCYVQVGLKPLVQECMVHDCADKLFQC 140
    :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  163 RHGIFTPETNPAGLEK-NYCRNPDPDGVNGPWCYT-TNPKRLFYCDIPQC---ESSFDC 217

QY  141 GOKTLRPR---PKIIGBFTTIENOPFAATYRRHRGGSVTYVCGGSLISPCWVISATHC 197
    :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  218 GKPKVEPKKCPARVVGGCVATPHSPWQVSLRRSR----BHFCCGTLISPEWVLTAAHC 273

QY  198 FIDYPKKEDYIVLGRSLNNTQGMKFVENLILHKDYSADTLAHNDIALIKIRSK 257
    :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  274 LDSILGPSFYVILGAHYEMAREASVQEIPIVSRJLFLEPSRA-----DIALKLGSP- 324

QY  258 GRCAQPSRTIOTICLPSMYNDPQF----GTSCETIGFGKENSTDYLPQEQKMTVWLIS 313
    :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  325 ---AVITDEVIPACLPSS---PNYVADKTVCYITGWGETQGT--FGVGRLEKRLPVIE 375

QY  314 HRECOQPHYGVSEVNTKMLCAADPOWKTDSCGGSGGLVCLSGLOGRMTLTIGVSWGRC 373
    :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  376 NKVCNRYEYLNCRVVKSTELCAGDLAGGTDCQSGSGGLVCFEKDKYILQGVTSWGLGCA 435

QY  374 LKDKPGVYTVRSHFLPWI 391
    :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  436 RPNKPGVYTVRSTVPWI 453
```

RESULT 24

PLHU
plasmin (EC 3.4.21.7) precursor [validated] - human
N:Alternate names: plasminogen precursor [misnomer]
N:Contains: angiotensin; microplasmin; plasminogen
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1984 #sequence revision 02-Dec-1994 #text change 15-Sep-2000
C:Accession: A35229; I52242; A26646; I62738; I84609; S03735; A00929; A04627; A04625; A04
R:Peterson, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.
J. Biol. Chem. 265, 6104-6111, 1990
A:Title: Characterization of the gene for human plasminogen, a key proenzyme in the fibr
A:Reference number: A35229; MUID:90202879; PMID:2318848
A:Accession: A35229
A:Molecule type: DNA
A:Residues: 1-810 <P>
A:Cross-references: GB:J05286; GB:M34276; NID:G190064; PIDN:AAA60113.1; PID:G387026
A:Experimental source: leukocyte; lung fibroblast
R:Margaretta, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Mironi, G.; Ottolenghi, S.; Ta
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990
A:Title: Definition of the transcription initiation site of human plasminogen gene in li
A:Reference number: I52242; MUID:91097523; PMID:2268308
A:Accession: I52242
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-16 <MALI>
A:Cross-references: GB:M62890; NID:G190092; PIDN:AAA36454.1; PID:G553613
R:Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
FEBS Lett. 213, 254-260, 1987
A:Title: Molecular cloning and characterization of a full-length cDNA clone for human pl
A:Reference number: A26646; MUID:87162490; PMID:3030813
A:Accession: A26646
A:Molecule type: mRNA
A:Residues: 1-471, 'D', 473-810 <FOR>
A:Cross-references: GB:X05199; NID:G35530; PIDN:CAA28831.1; PID:G35531

A:Experimental source: liver
R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A:Title: Characterization of a complementary deoxyribonucleic acid coding for human an
A:Reference number: I45961; MUID:85023311; PMID:6148961
A:Accession: I62738
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 232-471, 'D', 473-810 <MAL2>
A:Cross-references: GB:K02922; NID:G190112; PIDN:AAA60124.1; PID:G387031
A:Accession: I84609
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 367-419 <MAL3>
A:Cross-references: GB:K02921; NID:G190110; PIDN:AAA60123.1; PID:G190111
R:Brumsholtz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.
Eur. J. Biochem. 114, 465-470, 1981
A:Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
A:Reference number: S03735; MUID:81212097; PMID:7238497
A:Accession: S03735
A:Molecule type: protein
A:Residues: 20-71, 'E', 73-76 <BRU>
R:Sostrup-Jensen, L.; Petersen, T.E.; Magnusson, S.
submitted to the Atlas, July 1977
A:Reference number: A00929
A:Accession: A00929
A:Molecule type: protein
A:Residues: 20-71, 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>
R:Wiman, B.
Eur. J. Biochem. 76, 129-137, 1977
A:Title: Primary structure of the B-chain of human plasmin.
A:Reference number: A04627; MUID:77225245; PMID:142009
A:Accession: A04627
A:Molecule type: protein
A:Residues: 581-810 <WIL>
R:Wiman, B.; Wallen, P.
Eur. J. Biochem. 50, 489-494, 1975
A:Title: Structural relationship between "glutamic acid" and "lysine" forms of human p
A:Reference number: A04625; MUID:75093329; PMID:122932
A:Accession: A04625
A:Molecule type: protein
A:Residues: 20-50, 'Q', 51-71, 'E', 73-85, 87-100 <W12>
R:Wiman, B.; Wallen, P.
Eur. J. Biochem. 58, 539-547, 1975
A:Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen t
A:Reference number: A04626; MUID:76043692; PMID:126863
A:Accession: A04626
A:Molecule type: protein
A:Residues: 483-507, 'E', 509-604 <W13>
R:Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.
J. Biol. Chem. 248, 1631-1633, 1973
A:Title: The primary structure of human plasminogen. II. The histidine loop of human p
A:Reference number: A92125; MUID:73149248; PMID:4694729
A:Contents: annotation; active site
R:Groskopf, W.R.; Summaria, L.; Robbins, K.C.
J. Biol. Chem. 244, 3590-3597, 1969
A:Title: Studies on the active center of human plasmin. Partial amino acid sequence of
A:Reference number: A92048; MUID:69234739; PMID:4240117
A:Contents: annotation; active site
R:Trexler, M.; Valli, Z.; Paddy, L.
J. Biol. Chem. 257, 7401-7406, 1982
A:Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen;
A:Reference number: A92382; MUID:82213905; PMID:6919539
A:Contents: annotation; omega-aminocarboxylic acid binding sites
R:Valli, Z.; Paddy, L.
J. Biol. Chem. 259, 13690-13694, 1984
A:Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essenti
A:Reference number: A92458; MUID:85054794; PMID:6094526
A:Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site
R:Cao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehnle, S.; McCance, S.G.
J. Biol. Chem. 271, 29461-29467, 1996
A:Title: Kringle domains of human angiotensin. Characterization of the anti-proliferat

A:Molecule type: mRNA
A:Residues: 1-810 <LUI>
A:Cross-references: GB:J04697; NID:G342272; PIDN:AAA36901.1; PID:G342273
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:1-9/Domain: signal sequence #status predicted <SIG>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:377-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:581-803/Domain: trypsin homology <TRY>
F:581-803/Domain: trypsin homology <TRY>
F:49-73, 53-61, 103-181, 124-164, 152-176, 185-262, 188-316, 206-245, 234-257, 275-352, 296-335, 32
bonds: #status predicted
F:622, 665/Active site: His, Asp, Ser #status predicted

Query Match 22.2%; Score 501; DB 2; Length 810;
Best Local Similarity 34.4%; Pred. No. 4.7e-31;
Matches 140; Conservative 45; Mismatches 165; Indels 56; Gaps 16;

QY 10 NCDCLNGTGVSNKYFSNIHWCNCPKFKFGQ-----HCEIDKSKTCYEGNGH 56
DB 428 NPDAKGPWCFTTDPSPVRWEYCNLKKCSGTGSGVAAPPVVAQLPDAETPSEEDCMFGNGK 487
QY 57 FYRKASTDTRGRCLPWNATVLOQTYHAH-----SDALQLGLGKHNYCRNPD-NRRRP 111
DB 488 GYRGAATVGTGTCQEWAA-----QEPHSRIFTPTETPRAGLEK-NYCRNPDGSDVGGP 541
QY 112 WCYVQVGLKPLVQECMVHDCADGLKFKCQCKTLRPR---FKIIGFEFTTIENQPF 168
DB 542 WCYT-TNPKLYDFCDVPQCAAS--SFDCGKQVPEKPKGPRVVGCVAYPHSWP-QI 596
QY 169 YRHRGSGVTYVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGE 228
DB 597 SLRTRUG--MHFCGGTLISPEWVLTAAHCKLSRSPSFYKVLGAHR-----EVHLEPHV 649
QY 229 ENLILHKDYSADTLAHNNDIALKIRSKGRCQAQPSRTIQTCLPSMYNDPOF---GTS 284
DB 650 QETEVSKMFSEPARA---DIALKLSPP---AIITDKVIPACLPS-----PNYVADRT 698
QY 285 CEITGFGKENSVDYLPQQLKTVKVLISHRECCQPHYVSEVTTKMLCAADPQWKDSC 344
DB 699 CFITGGETQGT--YGAGLLKEARLPVIEKNVRYEFLNGTVKTELCAHLAGGTDSC 756
QY 345 QGDSGGPLVCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWI 391
DB 757 QGDSGGPLVCFEKDKYILQVTSWGLGCAAPKPKGVYVRSRFTWI 803

RESULT 26
PLMS
plasmin (EC 3.4.21.7) precursor - mouse
N:Contains: angiotensin; plasminogen
C:Species: Mus musculus (house mouse)
C:Date: 20-Sep-1991 #sequence revision 01-Nov-1996 #text_change 18-Jun-1999
C:Accession: A38514; S48202; S48203
R:Degeen, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.
Genomics 8, 49-61, 1990
A:Title: Characterization of the cDNA coding for mouse plasminogen and localization of b
A:Reference number: A38514; MUID:91184812; PMID:2081600
A:Accession: A38514
A:Molecule type: mRNA
A:Residues: 1-812 <DEG>
A:Cross-references: GB:J04766; NID:G200402; PIDN:AAA50168.1; PID:G200403
R:Rijmen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A:Title: Characterization of the murine plasma fibrinolytic system.
A:Reference number: S48202; MUID:95010076; PMID:7521120
A:Accession: S48202
A:Molecule type: protein
A:Residues: 20-25 <LUI>
A:Accession: S48203

A:Molecule type: protein
A:Residues: 22-27 <LUI>
C:Comment: Plasminogen is synthesized by the kidney and is present in plasma and many c
C:Comment: plasminogen is converted into plasmin by plasminogen activators, both plasmi
mediately after dissociation from the clot. In the presence of the inhibitor, the activ
e inhibitor, the activation involves also removal of the activation peptide.
C:Comment: Stromelysin 1 (see PIR:KCMS1) acts on plasminogen to produce angiotatin. 1
eful in treating solid tumors.
C:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen a
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolo
C:Keywords: angiotensin inhibitor; blood; duplication; fibrinolysis; glycoprotein; hy
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:1-9/Domain: signal sequence #status predicted <SIG>
F:20-812/Product: plasminogen #status predicted <PRO>
F:20-96/Domain: activation peptide #status predicted <APT>
F:79-466/Product: angiotensin #status predicted <AST>
F:97-581, 582-812/Product: plasmin #status predicted <MAT>
F:97-581/Domain: chain A #status predicted <ACH>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:377-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:582-812/Domain: chain B #status predicted <BCH>
F:582-803/Domain: trypsin homology <TRY>
F:49-73, 53-61, 103-181, 124-164, 152-176, 185-262, 188-316, 206-245, 234-257, 275-352, 296-335, 3
bonds: #status predicted
F:78-79/Cleavage site: Glu-Asn (stromelysin 1) #status predicted
F:136,308/Binding site: carboxylate (Asn) (covalent) #status predicted
F:466-467/Cleavage site: Thr-Val (stromelysin 1) #status predicted
F:581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental
F:624, 667, 762/Active site: His, Asp, Ser #status predicted

Query Match 22.2%; Score 501; DB 1; Length 812;
Best Local Similarity 32.3%; Pred. No. 4.8e-31;
Matches 133; Conservative 50; Mismatches 165; Indels 64; Gaps 15;

QY 10 NCDCLNGTGVSNKYFSNIHWCNCPK--KFGQHCET-----DKSKTCYEGNGH 56
DB 428 NPDAKGPWCFTTDPSPVRWEYCNLKKCSGTGSGVVELPTVSOBSPGSDSETDCMYGNGK 487
QY 57 FYRKASTDTRGRCLPWNATVLOQTYHAH-----RSDALQLGLGKHNYCRNPD- 106
DB 488 DYRGTAVTAAGTFCQGWAA-----QEPHSRIFTPTQTNPRADL-----EKNYCRNPDG 536
QY 107 NRRPWCYVQVGLKPLVQECMVHDCADGLKFKCQCKTLRPR---FKIIGFEFTTIENQ 163
DB 537 DVNGPWCYT-TNPKLYDYCDIPLCASAS-SFECGKQVPEKPKGPRVVGCVANPHSWP 594
QY 164 WFAAIYRHRGSGVTYVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGE 223
DB 595 WQISLRTFTFTG---QHFCGGTLIAPEWVLTAAHCKLSRSPSFYKVLGAHEEYIRGLDV 651
QY 224 MKPEVENLILHKOYSADTLAHNNDIALKIRSKGRCQAQPSRTIQTCLPSMYNDPOF-- 281
DB 652 QETSVAKLILE-----PNNDIALKLD-----SRPATITDKVIPACLPS-----PNYV 695
QY 282 --GTSCEITGFGKENSVDYLPQQLKTVKVLISHRECCQPHYVSEVTTKMLCAADPQW 339
DB 696 ADRTICVITGWGETQGT--FGAGRLKEAQLPVIEKNVRYEFLNGTVKTELCAHLAGG 753
QY 340 KTSQCGDSGGPLVCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWI 391
DB 754 GVDSCQDSGGPLVCFEKDKYILQVTSWGLGCAAPKPKGVYVRSRFTWDI 805

RESULT 27
PLBO
plasmin (EC 3.4.21.7) precursor - bovine
N:Alternate names: plasminogen

A:Note: apo(a) gene 1 (nomenclature of reference I52415)
A:Accession: I65286
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-16 <RE4>
A:Cross-references: GB:M66878; NID:gl78782; PIDN:AAA51749.1; PID:g553186
C:Genetics:
A:Gene: GDB:LPA
A:Cross-references: GDB:I20699; OMIM:152200
A:Map position: 6q26-q27
A:Note: several genes closely linked on chromosome 6 are identical in the first coding exons of kringle repeats
C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-4548/Product: apolipoprotein(a) #status experimental <MAT>
F:28-105/Domain: kringle homology <KR1>
F:142-219/Domain: kringle homology <KR2>
F:256-333/Domain: kringle homology <KR3>
F:370-447/Domain: kringle homology <KR4>
F:484-561/Domain: kringle homology <KR5>
F:598-675/Domain: kringle homology <KR6>
F:712-789/Domain: kringle homology <KR7>
F:826-903/Domain: kringle homology <KR8>
F:940-1017/Domain: kringle homology <KR9>
F:1054-1131/Domain: kringle homology <KR10>
F:1168-1245/Domain: kringle homology <KR11>
F:1282-1359/Domain: kringle homology <KR12>
F:1396-1473/Domain: kringle homology <KR13>
F:1510-1587/Domain: kringle homology <KR14>
F:1624-1701/Domain: kringle homology <KR15>
F:1738-1815/Domain: kringle homology <KR16>
F:1852-1929/Domain: kringle homology <KR17>
F:1966-2043/Domain: kringle homology <KR18>
F:2080-2157/Domain: kringle homology <KR19>
F:2194-2271/Domain: kringle homology <KR20>
F:2308-2385/Domain: kringle homology <KR21>
F:2422-2499/Domain: kringle homology <KR22>
F:2536-2613/Domain: kringle homology <KR23>
F:2650-2727/Domain: kringle homology <KR24>
F:2764-2841/Domain: kringle homology <KR25>
F:2878-2955/Domain: kringle homology <KR26>
F:2992-3069/Domain: kringle homology <KR27>
F:3106-3183/Domain: kringle homology <KR28>
F:3220-3297/Domain: kringle homology <KR29>
F:3334-3411/Domain: kringle homology <KR30>
F:3448-3525/Domain: kringle homology <KR31>
F:3562-3639/Domain: kringle homology <KR32>
F:3676-3753/Domain: kringle homology <KR33>
F:3782-3859/Domain: kringle homology <KR34>
F:3896-3973/Domain: kringle homology <KR35>
F:4010-4087/Domain: kringle homology <KR36>
F:4124-4201/Domain: kringle homology <KR37>
F:4228-4307/Domain: kringle homology <KR38>
F:4328-4541/Domain: trypsin homology <TRY>

Query Match 21.2%; Score 478; DB 1; Length 4548;
Best Local Similarity 36.0%; Pred. No. 1.9e-28;
Matches 129; Conservative 36; Mismatches 133; Indels 60; Gaps 17;

QY 50 CYEGNGHYRGKASTDTMGRCPLPWSATVQLQTYHAHRSDALQLGLGK-----HNYCRN 104
Db CMFNGKGYRGKATTVTGPCQEW-----AAQEPHRHSTFTPGTNKWAGLEKRYCRN 4280

QY 105 PD-NRRPWCYVQVGLKPLVQECMVHDCADGLKFCQCGKTLRPR---FKIIGGEFTTIE 160
Db PDGDIINGWCYT-MNPKLFYCDIPLCASS--SFDGKQVQEPKPCPSIVGGCVAPHR 4337

QY 161 NQWFAAIYRRHGGSVTYVCGGSLISPCWIGATHCFIDYPKKEDYIVYLGSR---LN 217
Db SWPQVSL--RTRFGK--HFCGGTLISPEWVLTAAHCLKKSRPSSYKVLGAHQEVNLE 4393

QY 218 SNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSEGRCAQPSRTIQTICLPSMYN 277

Db 4394 SHVQ---EIEVSRLEFLEPTQA-----DIALKLK-----SRPAVITDKVMACLPSS--- 4435
QY 278 DPQF-----QTSCEITGFGKENSTDYLPQLKMTVVVVKLISHRECCQPHYGGSEVTTQMLC 333
Db 4436 -PDYMTARTCEVITWGGETQGT--FGTGLLKEAQLLVIEVNCN--HY-----KYIC 4483
QY 334 AADPQWKTDSCGDSGGLVCSLQGRMTLTGIVSGRCALKDKGVVTRVSHFLPMI 391
Db 4484 AEHLARGTUSCGDSGGLVCFEKDKYILQGVTSWGLGCARPKNKPGVVARVSRFVTWI 4541

RESULT 29
A32869
A:apolipoprotein(a) (EC 3.4.21.-) - rhesus macaque (fragment)
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 22-Jun-1999
C:Accession: A32869; A30848
R:Tominson, J.E.; McLean, J.W.; Lawn, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A:Reference number: A32869; MUID:89174660; PMID:2925643
A:Accession: A32869
A:Molecule type: mRNA
A:Residues: 1-1420 <TOM>
A:Cross-references: GB:J04635; NID:G342072; PIDN:AAA36833.1; PID:G342073
C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F:50-127/Domain: kringle homology <KR1>
F:164-241/Domain: kringle homology <KR2>
F:278-355/Domain: kringle homology <KR3>
F:392-469/Domain: kringle homology <KR4>
F:506-583/Domain: kringle homology <KR5>
F:620-697/Domain: kringle homology <KR6>
F:726-803/Domain: kringle homology <KR7>
F:840-917/Domain: kringle homology <KR8>
F:954-1031/Domain: kringle homology <KR9>
F:1068-1145/Domain: kringle homology <KR10>
F:1191-1413/Domain: trypsin homology <TRY>

Query Match 21.1%; Score 477; DB 2; Length 1420;
Best Local Similarity 34.0%; Pred. No. 6.4e-29;
Matches 129; Conservative 43; Mismatches 137; Indels 70; Gaps 15;

QY 50 CYEGNGHYRGKASTDTMGRCPLPWSATVQLQ--QTYHAHRSDALQLGLGHXNYCRNPDN 107
Db 1068 CYHNGQSYRGTFSTVTGRTCCSWSSMTPHQKATPENHPNDLTM-----NYCRNPDA 1122
QY 108 RRRPWCYVQVGLKPLVQEC--CMVHDCAD-----GKLFQCGQ 142
Db 1123 DTGPNCF--MDPSVRREYCNLTCSDETEGTVTPPTVIVPVSLEAPSEQASSFDCGK 1179
QY 143 KTLRPR---FKIIGGEFTTIEQNPFAPAIYRRHGGSVTYVCGGSLISPCWIVISATHCFI 199
Db 1180 PQVEPKPCPSIVGGCVAPHRSPWQVSL--RTRFGK--HFCGGTLISPEWVLTAAACLE 1235
QY 200 DYPKKEDYIVYLGSR---LNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSK 256
Db 1236 TFSRSFTKVLGAHQEVNLESHVQ---EIEVSRLEFLEPGA-----DIALKLK--- 1281
QY 257 EGRCAQPSRTIQTICLPSMYNDPQF-----GTSCEITGFGKENSTDYLPQLKMTVVKLI 312
Db 1282 -SRPAITDKVAPCLPS---PNVITAWTECITWGGETQGT--FGAGLLKEAQLHVI 1334
QY 313 SHRECCQPHYGGSEVTTQMLCAAPQWKTDSCGDSGGLVCSLQGRMTLTGIVSWGRGC 372
Db 1335 ENTVCNHYEFLNGRVKSTELCAGHLAGGTDRCCQDNGGPPVCFDKKYLIRGITTSWGGC 1394
QY 373 ALKDKFGVYTRVSHFLPMI 391
Db 1395 ACPNKGYYVVRVSFVTWI 1413

```
RESULT 30
A61545
plasmin (EC 3.4.21.7) precursor - horse (fragments)
N:Alternate names: plasminogen
N:Contains: miniplasminogen
C:Species: Equus caballus (domestic horse)
C:Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
C:Accession: A61545; S17527
R:Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A:Title: Structural aspects of the plasminogen of various species.
A:Reference number: A61545; MUID:89005015; PMID:3168975
A:Accession: A61545
A:Molecule type: protein
A:Residues: 1-33;34-117 <SCH>
R:Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
Protein Seq. Data Anal. 4, 69-74, 1991
A:Title: Complete amino acid sequence of equine miniplasminogen.
A:Reference number: S17527; MUID:92052077; PMID:1946332
A:Accession: S17527
A:Molecule type: protein
A:Residues: 118-455 <SC2>
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase; 2
F:1-33/34-117/118-455/Product: plasminogen (fragments) #status experimental <PRO>
F:1-33/Domain: activation peptide (fragment) #status experimental <APT>
F:34-117/118-225/226-455/Product: plasmin (fragments) #status experimental <MAT>
F:37-114/Domain: kringle homology <KR4>
F:118-455/Product: miniplasminogen #status experimental <MIN>
F:126-205/Domain: kringle homology <KR5>
F:226-455/Domain: plasmin chain B #status experimental <BCH>
F:226-448/Domain: trypsin homology <TRY>
F:267,310,405/Active site: His, Asp, Ser #status predicted

Query Match      20.0%; Score 451.5; DB 2; Length 455;
Best Local Similarity 32.5%; Pred. No. 1.8e-27;
Matches 122; Conservative 45; Mismatches 169; Indels 39; Gaps 12;

QY 27 NIHWNCNPKKFGGQHCIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHA 86
Db 103 SVWNEFCNLKCSQETVQSEPDCEMLGIGYGGKATVTGTRCOAWAA-----QEPHR 157
QY 87 H---RSDALQLGLGKHNYCRNPD-NRRRRPWCYVQVGLKPLVQECMVHDCADGKLKFOCGQ 142
Db 158 HSIFTPEANPWANLEKNYCRNPDGVDNPGWCYT-MNPQKLFYCDVPQCESS--PFDCK 214
QY 143 KTLRPR---FKIIGGEFTTIENOPFAATYRRHGGSVTVVCGSLISPCWVISATHCFI 199
Db 215 PKVEPKKCSGRVGGCVIAHSPW--QISLSTRFG--RHFCGGTLISPEWVLTAAHCL 270
QY 200 DYPKKEDYIVYLGSRSLNSNTQEMKFEVENILHKDYSADTLAHNDIALKIKESGR 259
Db 271 RSSRPSTYKVLGTHLRLAAGAQQIDVSKLFLEPSRA-----DIALKLSS---- 318
QY 260 CAQPSSTIQTI---CLPSMYNDPQFTSCETITGFGKENSTDYLYPEQLKQTVVVKLISHRE 316
Db 319 ---PAITQNVIPACLPAPADYVYVWNAECFVTGWGE--TQDSSNAGVLKEAQLPVENKV 373
QY 317 CQOPHYVGSVEVTTKMLCAADPQWKTDSCQSDSGGLVCSLQGRMTLTGIVSWGRCAKLD 376
Db 374 CNRYEVLNGRVKSTELCAGHLVGGVDSCQSDSGGLVCFEKDKYTLQGVTSWGLCARPN 433
QY 377 KPGVYTRVSHFLPWI 391
Db 434 KPGVYVRVSSFINWI 448
```

Search completed: December 3, 2003, 14:44:18
Job time : 23.8448 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:53 ; Search time 12.9491 Seconds
(without alignments)
1463.563 Million cell updates/sec

Title: US-09-880-503-6

Perfect score: 2257

Sequence: 1 SNELHQVPSNCDCLNGTGV.....VSHFLPWRSHTKENGLAL 403

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2240	99.2	431	1	UROK HUMAN
2	2090	92.6	433	1	UROK PAPCY
3	1827.5	81.0	442	1	UROK PIG
4	1728	76.6	433	1	UROK BOVIN
5	1656.5	73.4	432	1	UROK RAT
6	1626.5	72.1	433	1	UROK MOUSE
7	997	44.2	434	1	UROK CHICK
8	868.5	38.5	477	1	URT2 DESRO
9	862.5	38.2	431	1	URT2 DESRO
10	858.5	38.0	562	1	TPA HUMAN
11	841.5	37.3	559	1	TPA RAT
12	840.5	37.2	477	1	URT1 DESRO
13	824.5	36.5	559	1	TPA MOUSE
14	814	36.1	566	1	TPA BOVIN
15	756	33.5	394	1	URT3 DESRO
16	726.5	32.2	655	1	HGFA HUMAN
17	716	31.7	653	1	HGFA MOUSE
18	700.5	31.0	603	1	FAI2 CAURO
19	681	30.2	615	1	FAI2 HUMAN
20	638	28.3	593	1	FAI2 BOVIN
21	508.5	22.5	790	1	PLMN PIG
22	505	22.4	812	1	PLMN MOUSE
23	501	22.2	810	1	PLMN NACMU
24	499	22.1	810	1	PLMN HUMAN
25	498.5	22.1	333	1	PLMN CANFA
26	493	21.8	343	1	PLMN SHEEP
27	484.5	21.5	812	1	PLMN BOVIN
28	478	21.2	4548	1	APOA HUMAN
29	477	21.1	1420	1	APOA NACMU
30	447.5	19.8	338	1	PLMN HORSE
31	447	19.8	875	1	NETR HUMAN
32	434	19.2	761	1	NETR MOUSE
33	431	19.1	810	1	PLMN_ERIEU

ALIGNMENTS

RESULT 1

UROK HUMAN
ID UROK HUMAN STANDARD; PRT; 431 AA.
AC P00749; Q15844; Q16618; Q969W6;
DT 21-JUL-1986 (Rel. 01, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
DE (U-plasminogen activator).
GN PLAU.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=85215647; PubMed=2987867;
RX Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;
RT "The human urokinase-plasminogen activator gene and its promoter.";
RL Nucleic Acids Res. 13:2759-2771(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
RA Steffens G.J., Heynaker H.L.;
RT "Cloning and expression of the gene for pro-urokinase in Escherichia coli.";
RL Biotechnology 3:923-929(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86056954; PubMed=2415429;
RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,
RA Nishida M., Suyama T.;
RT "Molecular cloning of cDNA coding for human prepro-urokinase.";
RL Gene 36:183-188(1985).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=85203359; PubMed=3885571;
RA Jacobs P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P.,
RA van Elsen A., Herzog A., Sollen A.;
RT "Molecular cloning, sequencing, and expression in Escherichia coli of human prepro-urokinase cDNA.";
RL DNA 4:139-146(1985).
RN [5]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX TISSUE=Lung;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

34 425.5 18.9 418 1 HATT_HUMAN O60235 homo sapien
35 417 18.5 436 1 HEP5_MOUSE Q35453 mus musculus
36 408.5 18.1 343 1 PSS8_HUMAN Q16651 homo sapien
37 407.5 18.1 638 1 KAL_MOUSE P26262 mus musculus
38 406.5 18.0 455 1 TMS5_MOUSE Q9UKR3 homo sapien
39 405 17.9 277 1 KLKD_HUMAN P05981 homo sapien
40 404.5 17.9 417 1 HEP5_HUMAN P03952 homo sapien
41 402 17.8 638 1 KAL_HUMAN P03952 homo sapien
42 401.5 17.8 855 1 ST14_HUMAN Q9Y5Y6 homo sapien
43 401.5 17.8 855 1 ST14_MOUSE P56677 mus musculus
44 401 17.8 416 1 HEP5_RAT Q05511 rattus norv
45 395.5 17.5 248 1 TRY3_CHICK Q90629 gallus gall

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A., Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [7].

RP SEQUENCE OF 66-431 FROM N.A.

RX MEDLINE=8427206; PubMed=6589620;

RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F., "Identification and primary sequence of an unspliced human urokinase poly(A)+ RNA."

RT Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731 (1984).

RL [8].

RN SEQUENCE OF 21-177.

RP MEDLINE=8305084; PubMed=6754569;

RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E., Flohe L., "The primary structure of high molecular mass urokinase from human urine. The complete amino acid sequence of the A chain."

RT Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165 (1982).

RL [9].

RN SEQUENCE OF 156-176 AND 179-224.

RX MEDLINE=83003608; PubMed=6749491;

RA Schaller J., Nick H., Rickli E.E., Gillesse D., Lergier W., Studer R.O., "Human low-molecular-weight urinary urokinase. Partial characterization and preliminary sequence data of the two polypeptide chains."

RT Eur. J. Biochem. 125:251-257 (1982).

RL [10].

RN SEQUENCE OF 158-410.

RP MEDLINE=8305099; PubMed=6754572;

RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L., "The complete amino acid sequence of low molecular mass urokinase from human urine."

RT Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058 (1982).

RL [11].

RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).

RP MEDLINE=96000858; PubMed=8591045;

RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D., Dobson C.M., Stuart D.I., Jones E.Y., "The crystal structure of the catalytic domain of human urokinase-type plasminogen activator."

RT Structure 3:681-691 (1995).

RL [12].

RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.

RP MEDLINE=20266327; PubMed=10805774;

RA Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G., Bede W., Magdolen V., Huber R., Moroder L., "4-aminomethylphenylguanidine derivatives as nonpeptidic highly selective inhibitors of human urokinase."

RT Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118 (2000).

RL [13].

RN STRUCTURE BY NMR.

RP MEDLINE=89127526; PubMed=2536903;

RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M., "Dynamics of the multidomain fibrinolytic protein urokinase from two-dimensional NMR."

RT Nature 337:579-582 (1989).

RL [14].

RN STRUCTURE BY NMR OF 67-155.

RP MEDLINE=93003110; PubMed=1327118;

RA Li X., Smith R.A.G., Dobson C.M., "Sequential 1H NMR assignments and secondary structure of the kringle domain from urokinase."

RT Biochemistry 31:9562-9571 (1992).

RL [15].

RN STRUCTURE BY NMR OF 67-155.

RX MEDLINE=94149701; PubMed=8107091;

RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M., "Solution structure of the kringle domain from urokinase-type plasminogen activator."

RT J. Mol. Biol. 235:1548-1559 (1994).

RL [16].

RN VARIANT LEU-141.

RP MEDLINE=96186279; PubMed=8652631;

RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K., Sawasaki Y., Hanada K., "Characterization of single chain urokinase-type plasminogen activator with a novel amino-acid substitution in the kringle structure."

RT Biochim. Biophys. Acta 1293:83-89 (1996).

RL [17].

RN VARIANT LEU-141.

RP MEDLINE=97218551; PubMed=9065988;

RA Conne B., Berczy M., Belin D., "Detection of polymorphisms in the human urokinase-type plasminogen activator gene."

RT Thromb. Haemost. 77:434-435 (1997).

RL [18].

RN ERRATUM.

RA Conne B., Berczy M., Belin D., "Thromb. Haemost. 78:973-973 (1997).

RL [19].

RN VARIANT LEU-141.

RP MEDLINE=97337920; PubMed=9194591;

RA Turkmen B., Schmitt M., Schmalzfeldt B., Trommler P., Hell W., Kreutzburg S., Graeff H., Magdolen V., "Mutational analysis of the genes encoding urokinase-type plasminogen activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer."

RT Electrophoresis 18:686-689 (1997).

RL CC - FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR THERAPY OF THROMBOLYTIC DISORDERS.

CC CC - CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.

CC CC - SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.

CC CC - PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used in Pulmonary Embolism (PE) to initiate fibrinolysis.

CC CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC CC - SIMILARITY: Contains 1 kringle domain.

CC CC - SIMILARITY: Contains 1 EGF-like domain.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; X02419; CAA26268.1; -

DR EMBL; M15476; AAA61253.1; -

DR EMBL; D00244; BAA00175.1; -

DR EMBL; D11143; BAA01919.1; -

DR EMBL; X02760; CAA26535.1; -

DR EMBL; AF377330; AAK53822.1; -

DR EMBL; BC013575; AAH13575.1; -

DR EMBL; K03226; AAC97138.1; -

DR EMBL; K02286; AAA61252.1; -

DR EMBL; A21571; CAA01559.1; -

DR EMBL; A18397; CAA01390.1; -

DR PIR; A00931; UKHU.

DR PDB; 1KDU; 31-OCT-93.

Query Match 99.2%; Score 2240; DB 1; Length 431;
Best Local Similarity 97.8%; Pred. No. 7.4e-171; Indels 8; Gaps 1;
Matches 402; Conservative 1; Mismatches 0

QY 1 SNEHQVPSNCDLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
DB 21 SNEHQVPSNCDLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 80
QY 61 KASDTMTGRPCLPNSATVLCQTHAHSALQGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 81 KASDTMTGRPCLPNSATVLCQTHAHSALQGLGKHNYCRPNDRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQGOKTLRPRFKIIGGFTTIENQFWFAAIYRRH 172
DB 141 PLVQECMVHDCADGKSPPEELKFCQGOKTLRPRFKIIGGFTTIENQFWFAAIYRRH 200
QY 173 RGSVTVVCGSLSPCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 232
DB 201 RGSVTVVCGSLSPCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALALKIRSGRCAQPSRTIOTICLPSMYNDPQFGTSCIEITGFGK 292
DB 261 LHKDYSADTLAHHNDIALALKIRSGRCAQPSRTIOTICLPSMYNDPQFGTSCIEITGFGK 320
QY 293 ENSTDYILPQLKVTWVVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPL 352
DB 321 ENSTDYILPQLKVTWVVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPL 380
QY 353 VCSLQGRMTLTGVSGRGKALDKPGVYTRVSHFLPWIRSHKENGAL 403
DB 381 VCSLQGRMTLTGVSGRGKALDKPGVYTRVSHFLPWIRSHKENGAL 431

RESULT 2
UROK PAPCY STANDARD; PRT; 433 AA.

AC P1627;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9556;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thoracic aorta;
RX MEDLINE=90287734; PubMed=2113276;
RA Au Y.P.T., Wang T.W., Clowes A.W.;
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
RT plasminogen activator";
RL Nucleic Acids Res. 18:3411-3411(1990).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.

CC EMBL; X51935; CAA36200.1;
CC FIR; S14687; UKBAY.
DR HSP; P00749; LLMW.
DR MEROPS; S01.231;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle_1.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; kringle_1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYP SP; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS0070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYP SIN_DOM; 1.
DR PROSITE; PS00134; TRYP SIN_HIS; 1.
DR PROSITE; PS00135; TRYP SIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 433
FT CHAIN 21 176
FT CHAIN 155 176
FT CHAIN 178 433
FT DOMAIN 26 62
FT DOMAIN 69 150
FT DOMAIN 151 177
FT DOMAIN 178 433
FT DISULFID 30 38
FT DISULFID 32 50
FT DISULFID 52 61
FT DISULFID 167 298
FT DISULFID 208 224
FT DISULFID 216 287
FT DISULFID 315 384
FT DISULFID 347 363
FT DISULFID 374 402
FT ACT_SITE 223 223
FT ACT_SITE 274 274
FT ACT_SITE 378 378
FT CARBOHYD 324 324
SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

Query Match 92.6%; Score 2090; DB 1; Length 433;
Best Local Similarity 90.6%; Pred. No. 6.2e-159; Indels 12; Gaps 3;
Matches 375; Conservative 17; Mismatches 10

QY 1 SNEHQVPSNCDLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
DB 21 SREL-QVPSDCGLNGGTGCMNKFSSIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 79
QY 61 KASDTMTGRPCLPNSATVLCQTHAHSALQGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 80 KASDTMTGRSCLAWSATVLCQTHAHSALQGLGKHNYCRPNDRRPPWCYVQVGLK 139
QY 121 PLVQECMVHDCADGK-----LKFCQGOKTLRPRFKIIGGFTTIENQFWFAAIYRRH 172
DB 140 QRVQECMVHNCADGKSPPEELQFCQGOKTLRPRFKIIGGFTTIENQFWFAAIYRRH 199
QY 173 RGSVTVVCGSLSPCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 232
DB 200 RGSVTVVCGSLSPCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 259
QY 233 LHKDYSADTLAHHNDIALALKIRSGRCAQPSRTIOTICLPSMYNDPQ---FQTSCEITG 289

```

Db 260 LHEDVSADTLAHNDIALKIRSEGRCAQPSRIQTICLPSMYNDNDPPFGSCITG 319
QY 290 FKENSTDYLYPEOLKMTVVKLISHRECCQPHYHGVSEVTTKMLCAADPQWKTDSCQDSG 349
Db 320 FKENSTDYLYPEOLKMTVVKLVSHQKQCPHYHGVSEVTTKMLCAADPQWKTDSCQDSG 379
QY 350 GPLVCSLOGRWTLTGIVSGRGCCALKKPKGVYTVSHFLPWIRSHTEENGLAL 403
Db 380 GPLVCSIOGHWTLTGIVSGRGCCALKKPKGVYTVSHFLPWIRSHTEENGLAL 433

RESULT 3
UROK_PIG
ID UROK_PIG STANDARD; PRT; 442 AA.
AC P04155;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLA0.
OS Sus scrofa (pig)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=6096832;
RA Nagamine Y.; Pearson D.; Altus M.S.; Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
RL Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RN REVISION TO 241.
RA Nagamine Y.;
RL Submitted (DEC 1986) to the PIR data bank.
CC -! CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -! SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -! SIMILARITY: Contains 1 kringle domain.
CC -! SIMILARITY: Contains 1 EGF-like domain.
CC
CC This SWISS-PROT entry is copyright: It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X01648; CAA25806.1; -;
CC EMBL; X02724; CAA26511.1; -;
CC PIR; A00932; UKPG.
CC MEROPS; S01.231; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF like.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Serine protease_Try.
CC Pfam; PF00051; Kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC ProDom; PD00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00200; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; FALSE_NEG.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS50070; KRINGLE_2; 1.
CC PROSITE; PS50240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.

```

```

DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
FT CHAIN 190 442 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 CONNECTING PEPTIDE.
FT DOMAIN 154 189 KRINGLE.
FT DOMAIN 190 442 SERINE PROTEASE.
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .).
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).
FT DISULFID 220 236 BY SIMILARITY.
FT DISULFID 228 299 BY SIMILARITY.
FT DISULFID 324 393 BY SIMILARITY.
FT DISULFID 356 372 BY SIMILARITY.
FT DISULFID 383 411 BY SIMILARITY.
FT ACT_SITE 235 235 CHARGE RELAY SYSTEM.
FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
FT CONFLICT 241 241 Q -> H (IN REF. 1; CAA25806).
FT CONFLICT 242 242 Q -> H (IN REF. 1; CAA26511).
FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
SQ SEQUENCE 442 AA; 49116 MW; EE32FCEFS0131EE CRC64;

Query Match 81.0%; Score 1827.5; DB 1; Length 442;
Best Local Similarity 78.1%; Pred. No. 4.7e-138;
Matches 329; Conservative 32; Mismatches 41; Indels 19; Gaps 2;

QY 1 SHELHOV--PSNCDCLNGTCVSNKYFESNIHWNCNPKKFGGHCIDKSKTCYEGNGHSE 58
Db 21 SHELHOESGASNCGLNGKCVSYKFSNIQCSCKPKFGGHCIDTSQTCEGNGHSE 80
QY 59 RKGASTDTMGRECLPWNSTVLIQQTTHAHRSDALQGLGKKNYCRNPDNRPRPCWYVQVG 118
Db 81 RKGANTNTGRRCLPWNSTVLIQQTTHAHRSDALQGLGKKNYCRNPDNRPRPCWYVQVG 140
QY 119 LKPLVQECMVHDCAGD-----KLFQCGQKTLPRPKIIGGETTIEN 161
Db 141 LKQLVQECMVPCSGESHRPAYDGNPFSTPEKVEFCGQKALPRFKIVGKSTTIEN 200
QY 162 QWFPAIYRHRHGGSVTVVCGSLSPCWVVSATHCFIDYPKKEDYIVYLGSRSLNSNQ 221
Db 201 QWFPAIYRHRHGGSVTVVCGSLSPCWVVSATHCFINQQKEDYIVYLGSRSLNSNQ 260
QY 222 GEMKFEVENLIHKDYSADTLAHNDIALKIRSEGRCAQPSRIQTICLPSMYNDNDPQF 281
Db 261 GEMKFEVEKLIHEDYSADSLAHNDIALKIRSEGRCAQPSRIQTICLPSMYNDNDPQF 320
QY 282 GTSCEITGFGKENSTDYLYPEOLKMTVVKLISHRECCQPHYHGVSEVTTKMLCAADPQWK 341
Db 321 GASCEIVFGKEDPSDYLYPEOLKMTVVKLISHRECCQPHYHGVSEVTTKMLCAADPQWK 380
QY 342 DSCQDSGGLVCSLQGRWTLTGIVSGRGCCALKKPKGVYTVSHFLPWIRSHTEENGL 401
Db 381 DSCQDSGGLVCSLQGRWTLTGIVSGRGCCALKKPKGVYTVSHFLPWIRSHTEENGL 440
QY 402 A 402
Db 441 A 441

RESULT 4
UROK_BOVIN
ID UROK_BOVIN STANDARD; PRT; 433 AA.
AC Q05589; Q28209;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```


DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 CN PLAU.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Aortic endothelium;
 RX MEDLINE=93216119; PubMed=8385052;
 RA Kretzschmar J., Haendler B., Kojima S., Rifkin D.B.,
 RA Schleuning W.-D.;
 RT "Bovine urokinase-type plasminogen activator and its receptor:
 RT cloning and induction by retinoic acid.";
 RL Gene 125:177-183(1993).
 RN [2]
 RP SEQUENCE OF 12-433 FROM N.A.
 RC TISSUE=Kidney;
 RA Ravn P., Berglund L., Petersen T.E.;
 RT "Cloning and characterization of the bovine plasminogen activators uPA
 RT and tPA.";
 RL Int. Dairy J. 5:605-617(1995).
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- INDUCTION: By retinoic acid.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L03546; AA51419.1; -;
 CC EMBL; X85801; CAA59796.1; -;
 CC PIR; JN0560; JN0560.
 CC HSP; P00749; 1LWM.
 CC MEROPS; S01.231; -;
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR008209; EGF-like.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser. protease_Try.
 CC Pfam; PF00051; Kringle; 1.
 CC PRINTS; PR00722; Chymotrypsin.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000395; Kringle; 1.
 CC SMART; SM00130; KR; 1.
 CC SMART; SM00020; Tryp_Spc; 1.
 CC PROSITE; PS00022; EGF_1; 1.
 CC PROSITE; PS01186; EGF_2; FALSE NEG.
 CC PROSITE; PS00021; KRINGLE_1; 1-
 CC PROSITE; PS00070; KRINGLE_2; 1.
 CC PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 CC Kringle; EGF-like domain; Signal; Zymogen.
 KW SIGNAL 1 20
 FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
 FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
 FT DOMAIN 29 65 EGF-LIKE.
 FT DOMAIN 72 153 KRINGLE.
 FT DOMAIN 154 180 CONNECTING PEPTIDE.
 FT DOMAIN 181 433 SERINE PROTEASE.
 FT DISULFID 33 41 BY SIMILARITY.

FT DISULFID 35 53 BY SIMILARITY.
 FT DISULFID 55 64 BY SIMILARITY.
 FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 211 227 BY SIMILARITY.
 FT DISULFID 219 290 BY SIMILARITY.
 FT DISULFID 315 384 BY SIMILARITY.
 FT DISULFID 347 363 BY SIMILARITY.
 FT DISULFID 374 402 BY SIMILARITY.
 FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CONFLICT 189 189 A -> T (IN REF. 2).
 SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;
 Query Match 76.6%; Score 1728; DB 1; Length 433;
 Best Local Similarity 73.8%; Pred. No. 3.8e-130;
 Matches 305; Conservative 44; Mismatches 54; Indels 10; Gaps 2;
 QY 1 SNEHQV--PSNCDCLNGGTCVSNKYFSNIHWNCPCPKFGQHCEDKSKTCYEGNGHFY 58
 DB 21 SNEVHKGESGECGCLNGGCKVYKYFSNIQRCSCKPKFGQHCEDTSTKYQGNHSHY 80
 QY 59 RGKASTDTMGRCPLPNSATVLTQTYVHAHRSALQLGLGKHNYCRNPDRRPPWCYVQVG 118
 DB 81 RGKANRDLGRPCLAWDSPVLLKMYHAHRSDAIQLGLGKHNYCRNPDNRPPWCYVQIG 140
 QY 119 LKPLVQECMVHDCADGKL-----KFCGOKTLRPREKIIIGGETTIENQWFAAIYR 170
 DB 141 LKQFVFCMVQDCSVGKSPSSPREKEFCGQKALRPRKIVGGQVTNNAENQWFAAIYR 200
 QY 171 RHGGSVTVVCGSLISPCWISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVEN 230
 DB 201 RHGGSSITVLCGSLISPCWVVSATHCFIDHPKKENIYVYLGSRSLNSDTRGEMQFEVEK 260
 QY 231 LILHKDYSADTLAHHNDIALALKIRSKGECAPQSTIQTICLPSMYNDPQFGTSCITGF 290
 DB 261 LILHEDYSAESLAHHNDIALALKIRTSRGCAQPSRSIQTLCPPEHDAHSRTRCEITGF 320
 QY 291 GKENSFDYLYPQLKMTVVVKLIASHRECCQPHYGVSEVTTKMLCAADPQKMTSCQGDSSG 350
 DB 321 GKENSFDYRYSDELKMTFVSLVSHVSCQPHYGVSEVTTKMLCAADPQKMTSCQGDSSG 380
 QY 351 PLVCSLQGRMTLTGIVSWRGCGALADKDKPOVYTRVSHFLPWIRSHYKEENGLAL 403
 DB 381 PLVCTIQGRLLTGTIVSWRGDCAMKYKPGVYTRVSKFLPWINTHTRGEINLVL 433
 RESULT 5
 UROK_RAT
 ID UROK_RAT STANDARD; PRT; 432 AA.
 AC P29598;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Fischer 344;
 RX MEDLINE=92233409; PubMed=1568219;
 RA Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,
 RA Kefford R.F.;
 RT "Transcriptional and posttranscriptional activation of urokinase
 RT plasminogen activator gene expression in metastatic tumor cells.";
 RL Cancer Res. 52:2489-2496(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;

RA Rabbani S.A.;
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH
 CC CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM
 CC CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE
 CC LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X63434; CAA45028.1; -;
 CC EMBL; X65651; CAA46601.1; -;
 CC PIR; S24604; S18932.
 CC HSSP; P00749; 1KDU.
 CC MEROPS; S01_231; -;
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR006209; EGF-like.
 CC InterPro; IPR006210; EGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser. protease_Try.
 CC Pfam; PF00051; kringle; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000395; Kringle; 1.
 CC SMART; SM00181; EGF; 1.
 CC SMART; SM00130; KR; 1.
 CC SMART; SM00020; Tryp_Spc; 1.
 CC PROSITE; PS00022; EGF_1; 1.
 CC PROSITE; PS01186; EGF_2; FALSE_NEG.
 CC PROSITE; PS00021; KRINGLE_1; 1.
 CC PROSITE; PS00070; KRINGLE_2; 1.
 CC PROSITE; PS02040; TRYPsin_DOM; 1.
 CC PROSITE; PS00134; TRYPsin_HIS; FALSE_NEG.
 CC PROSITE; PS00135; TRYPsin_SER; 1.
 CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 CC Kringle; EGF-like domain; Zymogen; Signal.
 CC SIGNAL 1
 CC CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 CC CHAIN 20 177 CHAIN A (BY SIMILARITY).
 CC CHAIN 156 177 SHORT A CHAIN (A1) (BY SIMILARITY).
 CC CHAIN 179 432 CHAIN B (BY SIMILARITY).
 CC DOMAIN 27 63 EGF-LIKE.
 CC DOMAIN 70 151 KRINGLE.
 CC DOMAIN 152 178 CONNECTING PEPTIDE.
 CC DOMAIN 179 432 SERINE PROTEASE.
 CC DISULFID 31 39 BY SIMILARITY.
 CC DISULFID 33 51 BY SIMILARITY.
 CC DISULFID 53 62 BY SIMILARITY.
 CC DISULFID 168 300 INTERCHAIN (BY SIMILARITY).
 CC DISULFID 210 226 BY SIMILARITY.
 CC DISULFID 218 289 BY SIMILARITY.
 CC DISULFID 314 383 BY SIMILARITY.
 CC DISULFID 373 401 BY SIMILARITY.
 CC ACT_SITE 225 225 CHARGE RELAY SYSTEM.
 CC ACT_SITE 276 276 CHARGE RELAY SYSTEM.
 CC ACT_SITE 377 377 CHARGE RELAY SYSTEM.
 CC CONFLICT 16 16 N -> H (IN REF. 2).
 CC CONFLICT 24 24 E -> G (IN REF. 2).
 CC CONFLICT 332 332 D -> N (IN REF. 2).
 CC SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;
 SQ

Query Match 73.4%; Score 1656.5; DB 1; Length 432;
 Best Local Similarity 72.2%; Pred. No. 1.8e-124; Indels 9; Gaps 2;
 Matches 291; Conservative 45; Mismatches 58;
 QY 9 SNDCCLNGGTCVSNKYFNSIHMCNCPKFGGQHCIDSKTCYEGNGHFYRGKASTDTMG 68
 DB 29 SNCGCQNGGVCSYKYFSSIRRCSPKPKFGEHCIDTSKTCYHGNGQSYRGKANTDTKG 88
 QY 69 RPLCPNNSATVLOQTYHAHRSALQLGCKHNYCRPNRRPWCYVQVGLPLVQECMV 128
 DB 89 RPLCLWNSPAVLQQTYNAHRSALSLGLGCKHNYCRPNRRPWCYVQVGLQKQFVQECMV 148
 QY 129 HDCADGKLLK-----FQCGQKTLRPRFKITGGFTTIENQFWFAAIYRRHRGGS-VTY 179
 DB 149 QDCSLSKSSSTVDQCGFCGQKALRPRFKIVGGFTTVENQFWFAAIYLNKNGSPSPF 208
 QY 180 VCGSLLSPCWISATHCFIDYPKXEDYIVYLGRLNSNTGEMKFEVENILHKDYSA 239
 DB 209 KCGSLLSPCWASATHCFVNPQKKEEYVYVYLGQSKRNSYNGEMKFEVEQLIHEDFSD 268
 QY 240 DTLAHNDIALLKIRSKGRCQAPSRITQITICLPSMYNDPQFGTSCETITGFGKENSTDYL 299
 DB 269 ETLAFANDIALLKIRISTGQCAQPSRTIQTICLPRFGDAPFGSDCEITGFGQESATDIF 328
 QY 300 YPEQLKMTYVVKLIHSHRECQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQR 359
 DB 329 YPKDLKMSVVKIISHQCKQPHYGYSEINYNKMLCAADPEWKTDSCSDGGGGLICNIDGR 388
 QY 360 MLTGTIVSWGRGALKDKPKGVYTRVSHLPWTRSHTKKEENGLA 402
 DB 389 PTLSGIVSWGSGCAEKNGKGVYTRVSYFLNWIQSHIGBENGLA 431

RESULT 6
 UROK MOUSE
 ID UROK MOUSE STANDARD; PRT; 433 AA.
 AC P06869;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85179474; PubMed=2985393;
 RA Belin D., Vassalli J.-D., Compepine C., Godeau F., Nagamine Y.,
 RA Reich E., Koehler H.P., Duvoisin R.M.;
 RT "Cloning, nucleotide sequencing and expression of cDNAs encoding
 RT mouse urokinase-type plasminogen activator.";
 RL Eur. J. Biochem. 148:225-232(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88163489; PubMed=2831940;
 RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;
 RT "The murine urokinase-type plasminogen activator gene.";
 RL Biochemistry 26:8270-8279(1987).
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
 CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
 CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE LOW
 CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -


```
DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
DR PROSITE; PSS0134; TRYPSIN_HIS; 1.
DR PROSITE; PSS0135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 434 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 171 CHAIN A (BY SIMILARITY).
FT CHAIN 173 434 CHAIN B (BY SIMILARITY).
FT DOMAIN 36 72 EGF-LIKE.
FT DOMAIN 79 158 KRINGLE.
FT DOMAIN 159 172 CONNECTING PEPTIDE.
FT DOMAIN 173 434 SERINE PROTEASE.
FT DISULFID 40 48 BY SIMILARITY.
FT DISULFID 42 60 BY SIMILARITY.
FT DISULFID 62 71 BY SIMILARITY.
FT DISULFID 162 296 INTERCHAIN (BY SIMILARITY).
FT DISULFID 202 218 BY SIMILARITY.
FT DISULFID 210 285 BY SIMILARITY.
FT DISULFID 310 379 BY SIMILARITY.
FT DISULFID 342 358 BY SIMILARITY.
FT DISULFID 369 397 BY SIMILARITY.
FT ACT_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 373 373 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 228 228 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 434 AA; 49400 MW; BD881048DD666A55 CRC64;

Query Match 44.2%; Score 957; DB 1; Length 434;
Best Local Similarity 47.8%; Pred. No. 4.9e-72;
Matches 187; Conservative 64; Mismatches 122; Indels 18; Gaps 7;

QY 11 CDCLNGGTCVSNKFSNIHWCNCPKFGQCEIDKSKTCVGNHFGYRGKASTDTMGRP 70
Db |||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
40 CQCLNGGTCITYRFFSQIKRCLCEPGYGLHCEIDTNSICVSGNGEDYRGWADP ---G 95
QY 71 CLPNSATVLO-QYHAHRSALQGLGKHNKNCNPNRRPWCYVQVGLKPLVQE--CM 127
Db |||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
96 CLYDHPSVIRWGDYHADLKNALQGLGKHNKNCNPNRRPWCYTK--RYSIQETPCS 153
QY 128 VHCADGKLFQCCQKTLPRFKIIGGEFTTIEQNPWFATYRRHGGSVTYCGGSLIS 187
Db |||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
154 TIE---KCECTCGRSFSKYFKIVGGSQAQVEVTPWLAGIFQNM-GTQFLCGSLID 208
QY 188 PCWISATHCID----YPKKEDIVILGRSLNSNTQGMKFEVENLILHKOYSADTLA 243
Db |||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
209 PCWLTAHCFTNPKQPNKSVKVFGLKSLTNDHEQVFWDEIISHPOFTDHTGG 268
QY 244 HNDIALIKIRSGKRCQAPRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQ 303
Db |||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
269 NDNDIALIRIETASQCAVESNYRTVCLPEKNLNDNTWCEIAGYKGNSTDIYYAQR 328
QY 304 LKMTVVKLIHRECCQPHYGSEVTTKMLCAADPQWKTDSCQSDSGGPLVCSLQGRMTLT 363
Db |||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
329 LMSATVNLISQDCCKNYDSTRVTDNMVAGDPLWETDACKGDSGGPMVCEHNGRMTLY 388
QY 364 GIVSGRCAKDKPGVYTVRSHFLPWIRSH 394
Db |||||: : : : : : : : : : : : : : : : : : : : : : : : : : :
389 GIVSGDGCCKNAPGVTVIRYLNWIDSN 419

RESULT 8
URT2 DESRO
ID _URT2 DESRO STANDARD; PRT; 477 AA.
AC P15638;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.66) (DSPA
DE alpha-2) (BAT-PA) (T-plasminogen activator).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
```



```
DR PROSITE; PSS0070; KRINGLE 2; 1.  
DR PROSITE; PSS0240; TRYPSIN_DOM; 1.  
DR PROSITE; PSS0134; TRYPSIN_HIS; 1.  
DR PROSITE; PSS0135; TRYPSIN_SER; 1.  
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
KW Kringle; EGF-like domain; Signal; Multigene family.  
FT SIGNAL 1 36 POTENTIAL  
FT CHAIN 37 431 SALIVARY PLASMINOGEN ACTIVATOR BETA.  
FT DOMAIN 37 75 EGF-LIKE.  
FT DOMAIN 82 163 KRINGLE.  
FT DOMAIN 179 431 SERINE PROTEASE.  
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT_SITE 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 41 62 BY SIMILARITY.  
FT DISULFID 46 53 BY SIMILARITY.  
FT DISULFID 65 74 BY SIMILARITY.  
FT DISULFID 92 163 BY SIMILARITY.  
FT DISULFID 103 145 BY SIMILARITY.  
FT DISULFID 134 158 BY SIMILARITY.  
FT DISULFID 168 299 BY SIMILARITY.  
FT DISULFID 211 227 BY SIMILARITY.  
FT DISULFID 219 288 BY SIMILARITY.  
FT DISULFID 313 388 BY SIMILARITY.  
FT DISULFID 345 361 BY SIMILARITY.  
FT DISULFID 378 406 BY SIMILARITY.  
FT CARBOHYD 139 139 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 352 352 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 431 AA; 48221 MW; 699B5675B162CBF CRC64;  
  
Query Match 38.2%; Score 862.5; DB 1; Length 431;  
Best Local Similarity 44.1%; Pred. No. 2.4e-61;  
Matches 175; Conservative 59; Mismatches 146; Indels 17; Gaps 8;  
  
QY 9 SNCDCLNGTGVSNKVFNIHWCNPKYFGGHOCHIDKSKTCYEGNGHFYEGKASTDTMG 68  
DB 42 SELRCFNCGTQWQARAFDFF-VCCQPKYTGKQEVDTHTACYDKDQGYRTGWTSES 100  
  
QY 69 RPLCPWNSATVLQYTHAHRSDALQGLGKKNYCRNPDRRRPWCYQVGLKPLVQECMV 128  
DB 101 AQCIWNNSNLLTRTYNGRRSDAITLGLGNHNYCRNPDRNSKPCWYKAFKIFLFCV 160  
  
QY 129 HDCADGLKFKOG-KTLRPRFKIGGFTIENQWPAALYRRHGG-SVYVCGGSLI 185  
DB 161 PVCS-----KATCGLRKYEPQLHSTGGFTDITSHPWQAALFAQNRSSGERFLCGGILI 216  
  
QY 187 SPCWVISATHCFID-YPKKEDYIVYLGSRSLNSNTQGMKPEVENLILHKDYSADTLAHH 245  
DB 217 SSCWVLTAAHCFQERYPPQHLRV-LGRTYRVKPKGEQTEVEKCIIEEFDDT--YN 273  
  
QY 246 NDIALLKIRSEGRCAQPSRIQICLPSMYNDPQFGTSCITGFGKENSTDYLYPQLK 305  
DB 274 NDIALQLKSGSPQCAQESDVRALICLPEANLQLPDWTECELSGYGKHKSSSPFYSEQLK 333  
  
QY 306 MTVVKLISHRECCQPHYYSVETVMTLCAADPQWKT-----DSCQGDGSGGLVCSLQGR 359  
DB 334 EGHVRLYSSRCTSKFLNKVTNNMLCAGTRSGEYIPNVHDACQDGSGLVCMNDNH 393  
  
QY 360 MTLGIVSWGRGKALKDKPGYVTRVSHLPWIRHTK 396  
DB 394 MTLGLIISWGVCGGEKDPGYVTKVNYLGIWRDMR 430  
  
RESULT 10  
ID TPA_HUMAN  
AC P00750; Q15103; STANDARD; PRT; 562 AA.  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)  
DE (t-PA) (t-plasminogen activator) (Alteplase) (Retelplase).  
GN PLAT.
```

```
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
CX NCBI_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Melanoma;  
RX MEDLINE=83115262; PubMed=6337343;  
RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,  
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,  
RA Goeddel D.V., Collen D.;  
RT "Cloning and expression of human tissue-type plasminogen activator  
RT cDNA in E. coli.";  
RL Nature 301:214-221(1983).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal lung;  
RX MEDLINE=88262579; PubMed=3133640;  
RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;  
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA  
RT from human fetal lung cells.";  
RL Nucleic Acids Res. 16:5695-5695(1988).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88054470; PubMed=2824147;  
RA Reddy V.B., Garramone A.J., Saseak H., Wei C.-M., Watkins P., Galli J.,  
RA Hsiung N.;  
RT "Expression of human uterine tissue-type plasminogen activator in  
RT mouse cells using BPV vectors.";  
RL DNA 6:461-472(1987).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86156143; PubMed=3009482;  
RA Friesner Degen S.J., Rajput B., Reich E.;  
RT "The human tissue plasminogen activator gene.";  
RL J. Biol. Chem. 261:6972-6985(1986).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84298137; PubMed=6089198;  
RA NY T., Elgh F., Lund B.;  
RT "The structure of the human tissue-type plasminogen activator gene:  
RT correlation of intron and exon structures to functional and  
RT structural domains.";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86284200; PubMed=3030401;  
RA Harris T.J., Patel T., Marston F.A., Little S., Emtage J.S.,  
RA Odenakker G., Voiclaert G., Rombauts W., Billiau A., Somer P.;  
RT "Cloning of cDNA coding for human tissue-type plasminogen activator  
RT and its expression in Escherichia coli.";  
RL Mol. Biol. Med. 3:279-292(1986).  
RN [7]  
RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
RC TISSUE=Umbilical vein;  
RX MEDLINE=90192123; PubMed=2107528;  
RA Siebert P.D., Fong K.;  
RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from  
RT human endothelial cells.";  
RL Nucleic Acids Res. 18:1086-1086(1990).  
RN [8]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
```



```
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 380 380 E -> K (IN REF. 1).
SQ SEQUENCE 559 AA; 62903 MW; 7DBD3809C1D1C921 CRC64;

Query Match 37.3%; Score 841.5; DB 1; Length 559;
Best Local Similarity 36.5%; Pred. No. 1.5e-59;
Matches 180; Conservative 65; Mismatches 143; Indels 105; Gaps 12;

QY 3 ELHVP-----SNCDLNGTGVNKFNSNIHWNCNPKKGGQCEIDKSKTCYEGNGHY 58
DB 74 QCHSVPRVRSCEPCFNGGTCQALYFDF-VQCPDGFVGKCDIDTRATCFEGGITY 132
QY 59 RGKASTDMRGPCLPKMSATVLOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYV-QV 117
DB 133 RGTWSTAENGACINWNSALSQKPYGARRPNAIKLGLGNHNYCRNPDRDKPWCYVFK 192
QY 118 GL-----KPLVQECMV----- 128
DB 193 GKYTFECSTPACPKGPTEDCYGKGYTYRGTHFTTSKASCLPWNMSMILIGKTYTAW 252
QY 129 -----HDCADGKLRFP-----CGQKTLR-PRFKIIGG 154
DB 253 NSQALGLGRHNYCRNPDDAKPCHVMKDKRLTWECYDMSPCSTCGLRQYKQFQRIKGG 312
QY 155 EFTIENQWFAAYI-RRHGGSVTYVCGSLISPCWWSATGCFIDYPKKEDYIVLGR 213
DB 313 LFTDITSPWQAAIFVKNRSPGRFLCGGVLSSCWVLSAAHCFVERPPPHLKVVLGR 372
QY 214 SRLNSQTGEMKFEVENLILHKYSADTLAHHNDIALLKRSKRGCAQPSRTIQTICLP 273
DB 373 TYRVVPEEETFFIEKYIVHKEFDDT--YNDNIALQLRSDDSCQAESSVGTACL 430
QY 274 SWYNDPQF-----GTSCEITGPKENSTDYLYPEQLKMTYVVKLSHRECQPHYGVSEVT 329
DB 431 ----DPVQLPDMTECELSGSGKGEASPPFSDRLKEARVRLYPSRCTSHQNFNTITS 486
QY 330 KMLCAADP-----QWKYDSCGDSGGLVCSLQGRMTLTIGVSWRGCAKDKRPGVYTRV 384
DB 487 NMLCAGDRTGNGQDVHDACQDSSGGLVCMIDKMTLLGLISWGLGCGQKDPGYITKV 546
QY 385 SHFLPWIRSHYKE 397
DB 547 TNYLNIQDNMKQ 559

RESULT 12
URT1 DESRO
ID URT1 DESRO STANDARD; PRT; 477 AA.
AC P98119;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.66) (DSFA
alpha-1).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1337019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RW [2]
```

```
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RX Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
RN [3]
RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RC TISSUE=Salivary gland;
RX MEDLINE=98022741; PubMed=9354616;
RA Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,
RA Schleuning W.D., Bode W.;
RT "Catalytic domain structure of vampire bat plasminogen activator: a
RT molecular paradigm for proteolysis without activation cleavage.";
RL Biochemistry 36:13483-13493(1997).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. POTENT THROMBOLYTIC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
CC THE PRESENCE OF FIBRIN 1.
CC -!- SUBUNIT: Monomer.
CC -!- DOMAIN: THE FIBRONECTIN TYPE-1 DOMAIN MEDIATES BINDING TO FIBRIN,
CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
CC STIMULATION OF ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type 1 domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M63987; AAA31591.1; -
DR EMBL; M63986; AAA31592.1; -
DR PIR; JS0597; JS0597
DR PDB; 1A5I; 23-MAR-99.
DR MEROPS; S01.232; -
DR GlycoSuiteDB; P98119; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FNI; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family; 3D-structure.
```


RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences";
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 CC MANY OTHER PHYSIOLOGICAL EVENTS.
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
 CC BOND.
 CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
 CC -1- PFM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
 CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
 CC ARG-308 CATALYZED BY PLASMIN. TISSUE KALLIKREIN OR FACTOR XA.
 CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
 CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 2 kringle domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ENBL; J03520; AAA40470.1; ..
 CC ENBL; BC011256; AAH1256.1; ..
 CC PIR; A29941; A29941.
 CC HSP; P00750; IA5H.
 CC MEROPS; S01.232; ..
 CC MGD; MGI:J7610; Plat.
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR062209; EGF-like.
 CC InterPro; IPR000083; Fibronctnl.
 CC InterPro; IPR06210; IEGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00008; EGF; 1.
 CC Pfam; PF00039; fnl; 1.
 CC Pfam; PF00051; kringle; 2.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000395; Kringle; 2.
 CC SMART; SM00181; EGF; 1.
 CC SMART; SM00058; fnl; 1.
 CC SMART; SM00130; KR; 2.
 CC SMART; SM00020; Tryp_Spc; 1.
 CC PROSITE; PS00022; EGF_1; 1.
 CC PROSITE; PS01186; EGF_2; 1.
 CC PROSITE; PS01253; FIBRONECTIN_1; 1.
 CC PROSITE; PS00021; KRINGLE_1; 2.
 CC PROSITE; PS00020; KRINGLE_2; 2.
 CC PROSITE; PS02440; TRYPsin_DOM; 1.
 CC PROSITE; PS00134; TRYPsin_HIS; 1.
 CC PROSITE; PS00135; TRYPsin_SER; 1.
 CC Plasma; Plasma; Hydroxylase; Serine protease; Glycoprotein;
 KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
 FT SIGNAL 1 17 PROBABLE.
 FT PROPEP 18 29
 FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
 FT CHAIN 309 559 CHAIN.
 FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
 FT CHAIN 36 78 CHAIN.
 FT DOMAIN 36 78 FIBRONECTIN TYPE-I.

FT	DOMAIN	79	117	EGF-LIKE.
FT	DOMAIN	124	285	KRINGLE 1.
FT	DOMAIN	213	294	KRINGLE 2.
FT	DOMAIN	309	559	SERINE PROTEASE.
FT	ACT SITE	355	355	CHARGE RELAY SYSTEM.
FT	ACT SITE	404	404	CHARGE RELAY SYSTEM.
FT	ACT SITE	510	510	CHARGE RELAY SYSTEM.
FT	DISULFID	38	68	BY SIMILARITY.
FT	DISULFID	66	75	BY SIMILARITY.
FT	DISULFID	83	94	BY SIMILARITY.
FT	DISULFID	88	105	BY SIMILARITY.
FT	DISULFID	107	116	BY SIMILARITY.
FT	DISULFID	124	205	BY SIMILARITY.
FT	DISULFID	145	187	BY SIMILARITY.
FT	DISULFID	176	200	BY SIMILARITY.
FT	DISULFID	213	294	BY SIMILARITY.
FT	DISULFID	234	276	BY SIMILARITY.
FT	DISULFID	265	289	BY SIMILARITY.
FT	DISULFID	297	428	INTERCHAIN (BY SIMILARITY).
FT	DISULFID	340	356	BY SIMILARITY.
FT	DISULFID	348	417	BY SIMILARITY.
FT	DISULFID	442	516	BY SIMILARITY.
FT	DISULFID	474	490	BY SIMILARITY.
FT	DISULFID	506	534	BY SIMILARITY.
FT	CARBOHYD	149	149	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	481	481	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CONFLICT	260	260	G -> A (IN REF. 1).
FT	CONFLICT	325	325	P -> A (IN REF. 1).
SQ	SEQUENCE	559 AA; 63122 MW; 8CCEE2BD894514D9 CRC64;		

Query Match 36.5%; Score 824.5; DB 1; Length 559;
 Best Local Similarity 35.5%; Pred. No. 3.4e-58;
 Matches 175; Conservative 63; Mismatches 150; Indels 105; Gaps 9;

QY	3	ELHQP	-----SNCCLNGCTCVSNKYFNSNIHCNCPKAFGGQHCIEDKSKTCYENGHPY 58
Db	74	QCHSVFVRSCSEPRCNGGTCCQALYFSDP	-VCQCPDGFVGKRCDDIDTRATCFEEQITY 132
QY	59	RGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQGLGKHNCRNPNRRRPMCVQVG 118	
Db	133	RGTWSTAESGAECINWSSVLSLKPYNARRPNAIKLGLGNHNYCRNPRDLKPMCYVFKA 192	
QY	119	LKPLVQECMVHDCADGKUK	----- 137
Db	193	GKYTEFCSTPACPKGKSEDCYGVGVYRGTHSLTTSQASCLPANSLVLMGKSYTAWRT 252	
QY	138	-----FOCG-QKTLRPFKILCG 154	
Db	253	NSQALGLGRHNYCRNPDDARPCHVWKDKLWHEYCDMSPCTCGLRQYKRPQFRKGG 312	
QY	155	EFTTIENQPFPAIY-RRHRGGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVVLGR 213	
Db	313	LYTDITSHPWQAPIFVKNKSPGPERFLCGGVLISSCWLSAAHCFLEFPFNHLKVLGR 372	
QY	214	SRLNSNTQGMKFEVENILHKDYSADTLAHNDIALLKIRSEKRCQAQSPRTQTICLP 273	
Db	373	TYRVVPGEEBQTFEIKYIVHEEFDDDT--YDNDIALQLRSQKCAQESSVGTACLP 430	
QY	274	SMYNDPQF-----GTSCEITGFKENSTDYLPQELKMTVVKLISHRECQCPHYVGSVTT 329	
Db	431	----DPNLQLPDWTECELSGVGKHEASSPFSDLEKAEHVLVLPSSRCTSCHLENKVTN 486	
QY	330	KMLCAAP-----QWKYTDSCQSGSGGGLVCSLQGRMTLTGIVSWGRCALKDKPQVTVY 384	
Db	487	NMLCAGDTRSGGNQDLHDACQSGSGGLVCMVINKQMTLTGIIISWGLCGQKQKQVGVYTKV 546	
QY	385	SHFLPMIRSHTK 397	
Db	547	TNYLDWHDNWKQ 559	

RESULT 14
 TPA_BOVIN

ID TPA_BOVIN STANDARD; PRT; 566 AA.

AC Q28198;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, last sequence update)

DT 28-FEB-2003 (Rel. 41, last annotation update)

DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)

DE (t-PA) (t-plasminogen activator).

GN PLAT.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RA Ravn P., Berglund L., Petersen T.E.;

RT "Cloning and characterization of the bovine plasminogen activators uPA

RT and tPA";

RL Int. Dairy J. 5:605-617(1995).

CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOMEN PLASMINOGEN

CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY

CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT

CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND

CC MANY OTHER PHYSIOLOGICAL EVENTS.

CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in

CC plasminogen to form plasmin.

CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE

CC BOND.

CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.

CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER

CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER

CC ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.

CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A

CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -!- SIMILARITY: Contains 1 EGF-like domain.

CC -!- SIMILARITY: Contains 1 fibronectin type I domain.

CC -!- SIMILARITY: Contains 2 kringle domains.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL; X85800; CAA59795.1; -

CC HESP; P00750; IRTF.

CC MEROPS; S01.232; -

CC InterPro; IPR001314; Chymotrypsin.

CC InterPro; IPR006209; EGF-like.

CC InterPro; IPR000083; Fibrincthl.

CC InterPro; IPR000083; Fibrincthl.

CC InterPro; IPR006210; EGF.

CC InterPro; IPR000001; Kringle.

CC InterPro; IPR001254; Ser_protease_Try.

CC Pfam; PF00008; EGF; 1.

CC Pfam; PF00039; fnl; 1.

CC Pfam; PF00051; kringle; 2.

CC Pfam; PF00089; trypsin; 1.

CC PRINTS; PR00722; CHYMOTRYPSIN.

CC PRINTS; PR00018; KRINGLE.

CC ProDom; PD000395; Kringle; 2.

CC SMART; SM00181; EGF; 1.

CC SMART; SM00058; fnl; 1.

CC SMART; SM00130; KR; 2.

CC SMART; SM00020; TRYD_SPC; 1.

CC PROSITE; PS00022; EGF_1; 1.

CC PROSITE; PS01186; EGF_2; 1.

CC PROSITE; PS01253; FIBRONECTIN 1; 1.

CC PROSITE; PS00021; KRINGLE 1; 1.

CC PROSITE; PS50070; KRINGLE_2; 2.

DR PROSITE; PS50240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;

KW Plasma; Kringle; EGF-like domain; Repeat; Signal.

FT SIGNAL 1 21

FT PROPEP 22 33

FT CHAIN 34 566

FT CHAIN 34 314

FT CHAIN 315 566

FT DOMAIN 40 82

FT DOMAIN 83 121

FT DOMAIN 128 209

FT DOMAIN 219 300

FT DOMAIN 315 566

FT ACT_SITE 361 361

FT ACT_SITE 410 410

FT ACT_SITE 517 517

FT DISULFID 42 72

FT DISULFID 70 79

FT DISULFID 87 98

FT DISULFID 92 109

FT DISULFID 111 120

FT DISULFID 128 209

FT DISULFID 149 191

FT DISULFID 180 204

FT DISULFID 219 300

FT DISULFID 240 282

FT DISULFID 271 295

FT DISULFID 303 434

FT DISULFID 346 362

FT DISULFID 354 423

FT DISULFID 448 523

FT DISULFID 480 496

FT DISULFID 513 541

FT CARBOHYD 153 153

FT CARBOHYD 487 487

SQ SEQUENCE 566 AA; 63701 MW; 2EB6BB4B32276C3 CRC64;

Query Match 36.1%; Score 814; DB 1; Length 566;

Best Local Similarity 35.8%; Pred. No. 2.4e-57;

Matches 176; Conservative 60; Mismatches 155; Indels 100; Gaps 9;

QY 3 ELHQP-SNCD---CLNGGTCVSNKYFSGNIHWCNPKYFGGQHCIDKSKTCYEGNGHFY 58

DB 78 QCHSVFVRSCEPFCWNGGTCQALYSSDF-VQCPGFGMGLCEIDATATCYKQGVAY 136

QY 59 RKASTDTMGRECLPWN SATVLQTYHAHRSALQLGLGKKNYCRNPNRRPWCYQVVG 118

DB 137 RGTWSTAESGAECAANNSSGLAMKPYSGRRPNARLGLGNHNYCRNPNQDQSKPCYVFEKA 196

QY 119 LKPLVQECQVHDCA----- 132

DB 197 GYISEFCSTACAKVAEEDGDCVTGNGLAYGTRSHYTSKASCLPWNVSFLTSKIYTAW 256

QY 133 -----DGKLFQ-----CG-QKTLRPRFXII 152

DB 257 KSNAPALGLGKHNCNPDGDAQPWCHWKDRLQTLWEYCDVPQCVCGLRQYKRPQFRIK 316

QY 153 GGEFTIENQPFAAIY-RRHGGSVTVCCGSLISPCWVISATHCFIDYFKKEDIYVL 211

DB 317 GGLFADITSHFQAALFVNRSPGERFLCGGILISSCWLSAAHCFQERYPPHHLKVFL 376

QY 212 GRSLNSNTQGMKFEVENLILHKDYSADTLAHHNDIALKIRSKRGRCQAPSRITQIC 271

DB 377 GRTVRLVPGEEQFVEVEKYIIHKEFDODT--YNDIALHLKSDSLTCAESASVATIC 434

QY 272 LPSMYNPFQFTSCETGFGKENSTDYLPQLKMTYVKKLISHRECOQPHYGVSEVTKM 331

DB 435 LPDASLQLPDWTECELSGYGKHSSPPFFSERLKEAHVRLYPSSRCTSQHLNFRVTNNM 494

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver, and Serum;
RX MEDLINE=93252878; PubMed=7683665;
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
RA Kitamura N.,
RT "Molecular cloning and sequence analysis of the cDNA for a human
RT serine protease responsible for activation of hepatocyte growth
RT factor. Structural similarity of the protease precursor to blood
RT coagulation factor XII.";
RL J. Biol. Chem. 268:10024-10028(1993).
RN [2]
RP SEQUENCE OF 40-655 FROM N.A.
RA Zhao S., Odell C.;
RL Submitted (FEB 1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY
CC CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
CC -!- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND.
CC -!- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
CC -!- TISSUE SPECIFICITY: LIVER.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D14012; BAA03113.1; -;
CC EMBL; Z69923; CAA93803.1; -;
CC PIR; A46688; A46688.
CC HSSP; P00763; LDPO.
CC MEROPS; S01.228; -;
CC Genew; HGNC:4894; HGFAC.
CC MIM; 604552; -;
CC GO; GO:0005576; C:extracellular; TAS.
CC GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR000083; Fibrinctnl.
CC InterPro; IPR000562; FN_type_II.
CC InterPro; IPR006210; HGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00039; fn1; 1.
CC Pfam; PF00040; fn2; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; Chymotrypsin.
CC PRINTS; PR00013; FNtypeII.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000995; FN Type II; 1.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00058; FN1; 1.
CC SMART; SM00059; FN2; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01253; FIBRONECTIN_1; 1.
CC PROSITE; PS00023; FIBRONECTIN_2; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS50070; KRINGLE 2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydroxylase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 30
FT PROPEP 31 372 CLEAVED IN ACTIVE FORM
FT CHAIN 373 407 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT
FT CHAIN 408 655 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG
FT CHAIN 108 148 FIBRONECTIN TYPE-II.
FT DOMAIN 160 198 EGF-LIKE 1.
FT DOMAIN 200 240 FIBRONECTIN TYPE-I.
FT DOMAIN 241 279 EGF-LIKE 2.
FT DOMAIN 286 367 KRINGLE.
FT DOMAIN 408 655 SERINE PROTEASE.
FT ACT_SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 497 497 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 598 598 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 108 133 BY SIMILARITY.
FT DISULFID 122 148 BY SIMILARITY.
FT DISULFID 164 175 BY SIMILARITY.
FT DISULFID 169 186 BY SIMILARITY.
FT DISULFID 188 197 BY SIMILARITY.
FT DISULFID 202 230 BY SIMILARITY.
FT DISULFID 228 237 BY SIMILARITY.
FT DISULFID 245 256 BY SIMILARITY.
FT DISULFID 250 267 BY SIMILARITY.
FT DISULFID 269 278 BY SIMILARITY.
FT DISULFID 286 367 BY SIMILARITY.
FT DISULFID 307 349 BY SIMILARITY.
FT DISULFID 338 362 BY SIMILARITY.
FT DISULFID 394 521 INTERCHAIN (BY SIMILARITY).
FT DISULFID 432 448 BY SIMILARITY.
FT DISULFID 440 510 BY SIMILARITY.
FT DISULFID 535 604 BY SIMILARITY.
FT DISULFID 567 583 BY SIMILARITY.
FT DISULFID 594 622 BY SIMILARITY.
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 492 492 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 546 546 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 644 644 R -> Q (IN REF. 2).
SQ SEQUENCE 655 AA; 70681 MW; 2CF7F1E1B862ED7 CRC64;
Query Match 32.2%; Score 726.5; DB 1; Length 655;
Best Local Similarity 37.2%; Pred. No. 2.5e-50;
Matches 158; Conservative 53; Mismatches 161; Indels 53; Gaps 9;
QY 5 HQVPSNCDLNGTGVCSNKFYSNIHW-----CNCPKFGGQHCCEIDKSKTCYEGNGH 56
DB 242 HTACLSPCLNGGTC-----HLIVATGTTVCACPPGFAGRLCNIEPDERCFLNGT 292
QY 57 FYRGKASTDTMGRCPCLPWSNATVLOOTYHRSALQLGLGKNYCRPNDRRRWCYV- 115
DB 293 GYRGVASTSASGLSCLAWNSLLYQELHVDVSGAALLGLGPHAYCRPNDRPCWYV 352
QY 116 -----QVGLKPLVQECMVHCADGKLKFCQGQK-----TLRPREKI 151
DB 353 KDSALSWEYCRLEACESLTVQLSPDLLATLPEPASPRQ--ACGRHKKKTFILPR--I 408
QY 152 IGGFEFTTIENOPWFAAIYRRHGGSVTVCGGSLISPCWVTSATHCFIDYPKKEDYIVYL 211
DB 409 IGGSSSLPGSHPLAAIY---IGDS---FCAGSLVHTCWVVSAACFSHSPRDSVSVVL 462
QY 212 GSRRLNSNTQCEMKFEVENLILHKDYSADTLAHHNDIALKIRSKGRCACPSRTIOTIC 271
DB 463 GQHFENRTDVTQTGIEKIPTLYSVFNPDSH-DLVLRILKKKKDRCATRSQFVDPIC 521
QY 272 LPSMYNDPQFGTSCIEITGFGKENSTDYLPQLKMTVVVKLISHRECQPHYGVSEVTTKM 331

```

Db 522 LPERGTFPAGHKQIAGWGHLDENVSGYSSLRREALVPLVADHKCSSPEVYGADISPNM 581
QY 332 LCAADPOKWTDSQGSGLVSLQGRMTLTGIVSWGRCAKLDKPGVYTRVSHLPWI 391
Db 582 LCAGYFCKSDACQSGGGLACFKNGVAYLYGIISWDCGRLKPKPGVYTRVANYVDWI 641
QY 392 RSHTK 396
Db 642 NDIR 646

RESULT 17
HGFA_MOUSE
ID HGFA_MOUSE STANDARD; PRT; 653 AA.
AC Q9R098; Q9JKV4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFAC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=BALB/c;
RC
RA Itoh H., Kataoka H., Koono H.;
RT "Mouse hepatocyte growth factor activator.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J.,
RA Yang J., Huan Y.;
RT "Activation of HGF by endogenous HGF activator is required for
RT metanephric kidney morphogenesis in vitro.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING
CC IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF099017; AAF02489.1; -
CC EMBL; AF224724; AAF34712.1; -
CC DR HSSP; P00763; IDPO.
CC DR MEROPS; S01.228; -.
CC DR MGD; MGI:1859281; Hgfac.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR006209; EGF_Like.
CC InterPro; IPR000083; Fibronctn1.
CC InterPro; IPR000562; FN_Type_II.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.

```

```

DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0013; FNTYPEII.
DR PRINTS; PRO0018; KRINGLE.
DR PRODOM; PD000395; FN_Type_II; 1.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29
FT PROPEP 30 369
FT CHAIN 370 405
FT CHAIN 406 653
FT DOMAIN 105 145
FT DOMAIN 157 195
FT DOMAIN 197 237
FT DOMAIN 238 276
FT DOMAIN 283 364
FT DOMAIN 406 653
FT ACT_SITE 445 445
FT ACT_SITE 495 495
FT ACT_SITE 596 596
FT DISULFID 105 130
FT DISULFID 119 145
FT DISULFID 161 172
FT DISULFID 166 183
FT DISULFID 185 194
FT DISULFID 199 227
FT DISULFID 225 234
FT DISULFID 242 253
FT DISULFID 247 264
FT DISULFID 266 275
FT DISULFID 283 364
FT DISULFID 304 346
FT DISULFID 335 359
FT DISULFID 392 519
FT DISULFID 430 446
FT DISULFID 438 508
FT DISULFID 533 602
FT DISULFID 565 581
FT DISULFID 592 620
FT CARBOHYD 39 39
FT CARBOHYD 47 47
FT CARBOHYD 63 63
FT CARBOHYD 287 287
FT CARBOHYD 466 466
FT CARBOHYD 544 544
FT CONFLICT 164 164
SQ SEQUENCE 653 AA; 70567 MW; 80B4B20255DF7FDC CRC64;

```

Query Match 31.7%; Score 716; DB 1; Length 653;
 Best Local Similarity 36.3%; Pred. No. 1; 7e-49;
 Matches 154; Conservative 58; Mismatches 162; Indels 50; Gaps 8;

OY 5 HQVPSNCDCLNGGTCVSNKYFSNIHW-----CNCPRKXFGGQHCEIDSKSKTCYEGNGH 56

Db 239 HTACLSPLNGTCTCTPLGVAGRCFNVPEHCFLNGT 289
Qy 57 FYRGKASTDTWRCPLPNSATVLOQYHARSALQGLCKHNYCRPNRRPWCYVQ 116
Db 290 EYRGVASTAAGSLCLAWNSDLLYQELHVDVAVALGLGPHAYCRNPKDPRWCYV 349
Qy 117 VGLXPLVQECMVHDC-----ADGKLKFOCGK-----TLRPREKII 152
Db 350 KDNALSWEYCLTACESLARVHSQTEPLAALPESAPAVRTCGKRHKKTFLPR--II 407
Qy 153 GGETTINQWPAIYARRHGGVYVCGGSLSPCWISATHCFIDYPKKEDIYVLG 212
Db 408 GSSSLPGSHWLAAY----IGNS----FCAGSLVHTCWWSAACFANSPPRDSITVVLG 461
Qy 213 RSLNSNTQGMKFEVENLIHKQYSADTLAHHNDIALLKRSKEGRCACPSRTIQICL 272
Db 462 QHFNRTDVTQTGIEKYVPYTLYSVFNPNH-DLVILRLKKGCACVRSQVQICL 520
Qy 273 PSMYNDPOFGTSCEITGFGKENSTDYLYPEOLKVTYVKL-SHRECQOPHYVGSVTTKML 332
Db 521 PEAGSSPFTGHKQIAGHGHNDENVSSYNSLLEALVPLVADHKCSSPEVVGADISPNML 580
Qy 333 CAADPQWKTDSCQDGGSLVCSLQGRMTLTGIVSGRGKALDKPGVYTRVSHFLPWIR 392
Db 581 CAGYFCKSDACQDGGSLVCKXNGVAYLYGLISWGDGCGRLNKGVYTRVANYVDWIN 640
Qy 393 SHTK 396
Db 641 DRIR 644

RESULT 18

FA12_CAVPO STANDARD; PR7; 603 AA.
AC Q04962;
AT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF) (Fragment).
GN F12.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.
RC TISSUE=Liver;
RX MEDLINE=93003367; PubMed=1390917;
RA Semba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,
RA Kambata T., Okabe H.;
RT "Primary structure of guinea-pig Hageman factor: sequence around the
RT cleavage site differs from the human molecule.";
RL Biochim. Biophys. Acta 1159:113-121(1992).
CC -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor
CC VII to form factor VIIa and factor XI to form factor Xla.
CC -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC FACTOR XII TO FORM KALLIKREIN. PREKALLIKREIN IS CLEAVED BY
CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X68615; CAA48600.1; -
DR PIR; S28941; S28941.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.211; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR000083; Fibronctnl.
DR InterPro; IPR000582; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
FT NON_TER 1
FT SIGNAL <1 18
FT CHAIN 19 358 ALPHA-FACTOR XIIA HEAVY CHAIN.
FT CHAIN 359 603 ALPHA-FACTOR XIIA LIGHT CHAIN.
FT DOMAIN 46 87 FIBRONECTIN TYPE-II.
FT DOMAIN 93 130 EGF-LIKE 1.
FT DOMAIN 132 172 FIBRONECTIN TYPE-I.
FT DOMAIN 173 209 EGF-LIKE 2.
FT DOMAIN 216 294 KRINGLE.
FT DOMAIN 312 342 PRO-RICH.
FT DOMAIN 359 603 SERINE PROTEASE.
FT ACT_SITE 398 398 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 551 551 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 97 109 BY SIMILARITY.
FT DISULFID 103 118 BY SIMILARITY.
FT DISULFID 120 129 BY SIMILARITY.
FT DISULFID 134 162 BY SIMILARITY.
FT DISULFID 160 169 BY SIMILARITY.
FT DISULFID 177 188 BY SIMILARITY.
FT DISULFID 182 197 BY SIMILARITY.
FT DISULFID 199 208 BY SIMILARITY.
FT DISULFID 216 294 BY SIMILARITY.
FT DISULFID 237 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 345 472 BY SIMILARITY.
FT DISULFID 383 399 BY SIMILARITY.
FT DISULFID 391 461 BY SIMILARITY.


```
FT DISULFID 422 425 BY SIMILARITY.
FT DISULFID 488 557 BY SIMILARITY.
FT DISULFID 520 536 BY SIMILARITY.
FT DISULFID 547 578 BY SIMILARITY.
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 603 AA; 66795 MW; 48DCB946FB9ED59 CRC64;

Query Match 31.08; Score 700.5; DB 1; Length 603;
Best Local Similarity 36.78; Pred. No. 2.7e-48;
Matches 160; Conservative 60; Mismatches 147; Indels 69; Gaps 13;

Qy 13 CLNGCTCVSNKYFNSIHWNCPCPKFGQGHCEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
Db 182 CLNGRCLE---VEGHLLCDPCMGYTGPFCDLDTTASCYEGRGVYRGVARTTVSGAKCQ 238
Qy 73 PWSATVLOQTYAHRSQ-ALQLGLGHNYCRPNDRRRRWCVVQVGLKPLVQECMWHDC 131
Db 239 RWAS-----EATYRNTAEQALRRGLGHHTFCRPNDRNDTRPWCFFVMGNRLSWCYCDLAQC 294
Qy 132 -----ADGKLKFP-----CGOKTLRPRF 149
Db 295 QYPPQPTATPHDRFEHFKLPSSRLSILQPTQONQALANELPETSSLLCGOR-LKRL 353
Qy 150 ----KIIGBFTTIENQPFAYRRHRRGGSVYVCGSLISPCWISATHCFIDYPKKE 205
Db 354 SLSRIVGLVALPGAHPHYAALY----WGS--NFCSGSLIAPCWVLTAAHCLQNRPAPE 407
Qy 206 DYIVYLCGRSLNSNTQGMKFVENLILHKDYADTLAHNDIALLKI-RSKSGRCAOPS 264
Db 408 ELKVVLQDRHQSCCHQTLVHSLHFAFSPS--SYLNDLALLRLQKSADGSCAQLS 465
Qy 265 RTIQTICLPSMYNDPQFG--TSCEITGFGKENSTDYLYPEQLXMTVVYKLISHRECOOPHY 322
Db 466 PYVQTVCLPSGAPPSSETTCCEVAGWGHQFEAGYEYSFLQBAQVPLISSERCSSPEV 525
Qy 323 YGSEVTTKMLCAADPQWTKSCQDSCGPLVC---SLQGRMTLTIGVSWRGCAKDKPG 379
Db 526 HGAFLGMLCAGLEGGTACQDGGPLVCEDEAAEHLLIRGIVSWGSGCDNRKPG 585
Qy 380 VYTRVSHFLPWIRSH 395
Db 586 VYTDVASYLWIOGHT 601

RESULT 19
FA12 HUMAN STANDARD; PRT; 615 AA.
AC P00748; P78339;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF).
GN F12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=88007593; PubMed=2888762;
RA Cool D.E., McGillivray R.T.A.;
RT "Characterization of the human blood coagulation factor XII gene.
RT Intron/exon gene organization and analysis of the 5'-flanking
RT region."
RL J. Biol. Chem. 262:13662-13673(1987).
[2]
RN SEQUENCE FROM N.A., AND VARIANTS ALA-207; ASP-545 AND HIS-605.
RP Rieder M.J., Arnel T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RN Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
```

```
RN SEQUENCE OF 4-615 FROM N.A.
RP MEDLINE=86176754; PubMed=3754331;
RX Tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A.,
RA Cortese R.;
RT "cDNA sequence coding for human coagulation factor XII (Hageman).";
RL Nucleic Acids Res. 14:3146-3146(1986).
[4]
RN SEQUENCE OF 14-615 FROM N.A.
RX MEDLINE=86033830; PubMed=3877053;
RA Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
RA McGillivray R.T.A.;
RT "Characterization of human blood coagulation factor XII cDNA.
RT Characterization of the primary structure of factor XII and the tertiary
RT structure of beta-factor XIIa."
RL J. Biol. Chem. 260:13666-13676(1985).
[5]
RN SEQUENCE OF 146-615 FROM N.A.
RP MEDLINE=86216049; PubMed=3011063;
RX Que B.G., Davie E.W.;
RA "Characterization of a cDNA coding for human factor XII (Hageman
RT factor).";
RL Biochemistry 25:1525-1528(1986).
[6]
RN SEQUENCE OF 20-379.
RP MEDLINE=85182674; PubMed=3886654;
RX McMullen B.A., Fujikawa K.;
RA "Amino acid sequence of the heavy chain of human alpha-factor XIIa
RT (activated Hageman factor).";
RL J. Biol. Chem. 260:5328-5341(1985).
[7]
RN SEQUENCE OF 354-362 AND 373-615.
RX MEDLINE=83291041; PubMed=6604055;
RA Fujikawa K., McMullen B.A.;
RT "Amino acid sequence of human beta-factor XIIa."
RL J. Biol. Chem. 258:10924-10933(1983).
[8]
RN SEQUENCE OF 561-615 FROM N.A.
RP TISSUE=Blood;
RX MEDLINE=96133302; PubMed=8528215;
RA Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammle B., Engel W.;
RT "The novel acceptor splice site mutation 11396(G-->A) in the factor
RT XII gene causes a truncated transcript in cross-reacting material
RL negative patients."
RL Hum. Mol. Genet. 4:1235-1237(1995).
[9]
RN CARBOHYDRATE-LINKAGE SITE THR-109.
RP MEDLINE=92184750; PubMed=1544894;
RX Harris R.J., Ling V.T., Spellman M.W.;
RT "O-linked fucose is present in the first epidermal growth factor
RT domain of factor XII but not protein C."
RL J. Biol. Chem. 267:5102-5107(1992).
[10]
RN VARIANT WASHINGTON D.C. SER-590.
RP MEDLINE=90046788; PubMed=2510163;
RX Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B.,
RA Saico H.;
RT "Coagulation factor XII (Hageman factor) Washington D.C.: inactive
RT factor XIIa results from Cys-571-->Ser substitution."
RL Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322(1989).
[11]
RN VARIANT LOCARNO PRO-372.
RX MEDLINE=94325559; PubMed=8049433;
RA Hovinga J.K., Schaller J., Stricker H., Wullemijn W.A., Furlan M.,
RA Lammle B.;
RT "Coagulation factor XII Locarno: the functional defect is caused by
RT the amino acid substitution Arg-353-->Pro leading to loss of a
RT kallikrein cleavage site."
RL Blood 84:1173-1181(1994).
[12]
RN VARIANT TENRI CYS-53.
RP MEDLINE=99290785; PubMed=10361128;
RX Kondo S., Tokunaga F., Kawano S., Oono Y., Kumagai S., Koide T.;
```

RT "Factor XII Tenri, a novel cross-reacting material negative factor XII
RT deficiency, occurs through a proteasome-mediated degradation.";
RL Blood 93:4300-4308 (1999).
CC -|- INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC -|- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-ile bonds in factor
CC VII to form factor VIIa and factor XI to form factor Xla.
CC -|- PPM: O- AND N-GLYCOSYLATED.
CC -|- DISEASE: DEFECTS IN F12 DO NOT CAUSE ANY CLINICAL SYMPTOMS. THE
CC SOLE EFFECT IS THAT WHOLE-BLOOD CLOTTING TIME IS PROLONGED.
CC -|- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XI FIRST
CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- SIMILARITY: Contains 2 EGF-like domains.
CC -|- SIMILARITY: Contains 1 fibronectin type I domain.
CC -|- SIMILARITY: Contains 1 fibronectin type II domain.
CC -|- SIMILARITY: Contains 1 kringle domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M1315; AAA70225.1; --
CC EMBL; AF538691; AAM57932.1; --
CC EMBL; M11723; AAS1986.1; --
CC EMBL; M17466; AAB59490.1; --
CC EMBL; M17464; AAB59490.1; JOINED.
CC EMBL; M17465; AAB59490.1; JOINED.
CC EMBL; M13147; AAA70224.1; --
CC EMBL; U71274; AAB51203.1; --
CC PIR; A29411; KFHUL2.
CC HSP; P00763; IDPO.
CC MEROPS; S01.211; --
CC MIM; 234000; --
CC GO; GO:0003805; F: blood coagulation factor XI activity; TAS.
CC GO; GO:0003806; F: blood coagulation factor XII activity; TAS.
CC GO; GO:0007595; P: blood coagulation; TAS.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR000083; Fibrinctn1.
CC InterPro; IPR000562; FN_Type_II.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00039; fn1; 1.
CC Pfam; PF00040; fn2; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00013; FNTYPEII.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000995; FN_Type_II; 1.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00058; FN1; 1.
CC SMART; SM00059; FN2; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01253; FIBRONECTIN_1; 1.
CC PROSITE; PS00023; FIBRONECTIN_2; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW Hydroxylase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;
KW Polymorphism; Disease mutation.
FT SIGNAL 1 19
FT CHAIN 20 372 ALPHA-FACTOR XIIA HEAVY CHAIN.
FT CHAIN 373 615 ALPHA-FACTOR XIIA LIGHT CHAIN.
FT CHAIN 354 362 BETA-FACTOR XIIA PART 1.
FT CHAIN 373 615 BETA-FACTOR XIIA PART 2.
FT DOMAIN 47 88 FIBRONECTIN TYPE-II.
FT DOMAIN 94 131 EGF-LIKE 1.
FT DOMAIN 133 173 FIBRONECTIN TYPE-I.
FT DOMAIN 174 210 EGF-LIKE 2.
FT DOMAIN 217 295 KRINGLE.
FT DOMAIN 296 349 PRO-RICH.
FT DOMAIN 373 615 SERINE PROTEASE.
FT CARBOHYD 109 109 O-LINKED (FUC).
FT CARBOHYD 249 249 N-LINKED (GLCNAC...).
FT CARBOHYD 299 299 O-LINKED (POTENTIAL).
FT CARBOHYD 305 305 O-LINKED (POTENTIAL).
FT CARBOHYD 308 308 O-LINKED (POTENTIAL).
FT CARBOHYD 328 328 O-LINKED (POTENTIAL).
FT CARBOHYD 329 329 O-LINKED (POTENTIAL).
FT CARBOHYD 337 337 O-LINKED (POTENTIAL).
FT ACT_SITE 412 412 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 461 461 CHARGE RELAY SYSTEM (BY SIMILARITY).

Query Match 30.2%; Score 681; DB 1; Length 615;
Best Local Similarity 34.1%; Pred. No. 9.7e-47;
Matches 152; Conservative 58; Mismatches 158; Indels 78; Gaps 10;

QY 13 CLNGGTCVSNKVFYNIHWNCNPKFGGQHCEDKSKTCYEGNGHGFYRGASDTDMGRPL 72
DB 183 CLHGGRCLE---VEGRLCHCPGVYTGPCDVTTRASCYDGRGLSYRGLARTITLSCAPQ 239
QY 73 PMSATVLQOQTY--HAHRSDALQGLGKHNYCNPNRRPWCYVQVGLKPLVQECMHDC 131
DB 240 PWAS----EATYRNVTAEQARNWGLGHAFCRPNPDNDRPWCFLVNRDRLSWEYCDLAQC 295
QY 132 -----ADGKL----- 136
DB 296 QTQTAAPPTVSPRLHVLMPAPAPPPQPTTRTPPOSQTFGALPAKREOPPSTLRNG 355
QY 137 KFQCGQ---KTLRPRFKIIGGEFTTIENQFWFAAIYRHRGSGSVTVVCGSLISPCWVIS 193
DB 356 PLSCGQRLKSLSSMTRVVGVLVALRGAHPYIAALYWGHS-----FCAGSLIAPCWILT 409
QY 194 ATHCFIDYPKEDYIYVLGSRSLNSNTQGMKEFEVENLILHKDYSADTLAHHNDIALKKI 253
DB 410 AAHCLQDRPAPELTVVLGQERNHSCPCQTLAVRSYRLHEAFS--PVSQYHDLALLRL 467
QY 254 R-SKEGRCAQPSRTIQTICLPSMYNDPQFTSCITGFGKENSTDYLYPEQLKMTVVVKLI 312
DB 468 QEDADGSCALLSPYVQPVCLPSGAARPESETTLQVAGWGHPGEGAEVASFLQEAQVPFL 527
QY 313 SHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGLVCSLQG---RMTLTGIVSWG 369
DB 528 SLERCSAPDVHSGSILPGLMCAFLGEGTDACQDGGGLVCEQDAERLTLQGIISWG 587
QY 370 RGCALKDKPGVYTRVSHFLPWIRSH 395
DB 588 SGCGRNKKPGVYTDVAYLLAWIREHT 613

RESULT 20
ID FA12_BOVIN STANDARD; PRT; 593 AA.
AC P98140;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 28-FEB-2003 (Rel. 41, last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF) (Fragment);
GN F12.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Liver;
RX MEDLINE=9424782; PubMed=8186251;
RA Shibuya Y., Samba U., Okabe H., Kambara T., Yamamoto T.;
RT "Primary structure of bovine Hageman factor (blood coagulation factor
RL Biochim. Biophys. Acta 1206:63-70(1994).
RN [2]
RN SEQUENCE OF 10-21; 350-364 AND 525-550.
RX MEDLINE=77182112; PubMed=861210;
RA Fujikawa K., Walsh A.K., Davie W.E.;
RT "Isolation and characterization of bovine factor XII (Hageman
RL factor).";
RN Biochemistry 16:2270-2278(1977).
CC -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-ile bonds in factor
CC VII to form factor VIIa and factor XI to form factor Xla.
CC -!- PTM: O- AND N-GLYCOSYLATED (BY SIMILARITY).
CC -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA. BOVINE FACTOR XII IS
CC CLEAVED ONLY TO ALPHA-FACTOR XIIA AS IT LACKS THE TRYPSIN/
CC KALLIKREIN CLEAVAGE SITE.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S70164; AB30804.2; -;
CC F1R; S45281; S45281.
CC HSP; P00763; IDPO.
CC MEROPS; S01.211; -;
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF like.
CC InterPro; IPR000083; Fibinctnl.
CC InterPro; IPR000562; FN Type_II.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00039; fn1; 1.
CC Pfam; PF00040; fn2; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00013; FNTYPEII.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000995; FN Type_II; 1.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 2.

DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR PROSITE; PS01253; FIBRONECTIN_I; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
FT NON_TER 1 1
FT SIGNAL <1 9 POTENTIAL.
FT CHAIN 10 349 ALPHA-FACTOR XIIA HEAVY CHAIN.
FT CHAIN 350 593 ALPHA-FACTOR XIIA LIGHT CHAIN.
FT DOMAIN 37 78 FIBRONECTIN TYPE-II.
FT DOMAIN 84 121 EGF-LIKE 1.
FT DOMAIN 123 163 FIBRONECTIN TYPE-I.
FT DOMAIN 164 200 EGF-LIKE 2.
FT DOMAIN 207 287 KRINGLE.
FT DOMAIN 297 333 PRO-RICH.
FT DOMAIN 350 593 SERINE PROTEASE.
FT ACT_SITE 389 389 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 438 438 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 541 541 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 88 100 BY SIMILARITY.
FT DISULFID 94 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 125 153 BY SIMILARITY.
FT DISULFID 151 160 BY SIMILARITY.
FT DISULFID 168 179 BY SIMILARITY.
FT DISULFID 173 188 BY SIMILARITY.
FT DISULFID 190 199 BY SIMILARITY.
FT DISULFID 207 287 BY SIMILARITY.
FT DISULFID 230 269 BY SIMILARITY.
FT DISULFID 258 282 BY SIMILARITY.
FT DISULFID 336 463 BY SIMILARITY.
FT DISULFID 374 390 BY SIMILARITY.
FT DISULFID 382 452 BY SIMILARITY.
FT DISULFID 413 416 BY SIMILARITY.
FT DISULFID 479 547 BY SIMILARITY.
FT DISULFID 510 526 BY SIMILARITY.
FT DISULFID 537 568 BY SIMILARITY.
FT CARBOHYD 99 99 O-LINKED (FUC) (BY SIMILARITY).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 593 AA; 65148 MW; 721592BA792BD61F CRC64;
Query Match 28.3%; Score 638; DB 1; Length 593;
Best Local Similarity 34.2%; Pred. No. 2.4e-43;
Matches 151; Conservative 59; Mismatches 164; Indels 68; Gaps 13;
QY 6 QVPSNCDCLNGGTCVSNKYFSNIHWCNPKKGGQHCIEDKSKTCYE--GNGHFYEGKAS 63
DB 166 QVCTNPCLNGDSCLOAE---GHLRCRCAPSFAGRLCDVLDKASCYDDDRDRGLSYEGMAG 222
QY 64 TDTMGSPCLPWN SATVLQTY-HAHRSDALQLGLGKHNYCRNPDRRRRWCVYQVGLKPL 122
DB 223 TTLSGAPCQSWAS-----EATYNNVTAEVLNWLGLGDHACRNPNDNTRPACFIWKDRLS 278
QY 123 VQECMVHDC--ADGKLFKQ-----CGQ 142
DB 279 WNYCLAPCQAAAGHEHFLPSPALQKPESTTQTPLSLTSGWCSPFTPLASGPGCGCQ 338
QY 143 ---KTLRPRFKIIGBEFTTIENQPFATYRRHGGSVTVVCGSLISPCWISATHCFI 199
DB 339 RLKRWLSSLNRRVVGGLVALPGAHPIYALYWDQ-----HFCAGSLIAPCWVLTAAHCLQ 392

QY 200 DYPKEDYIVYVGRSLNSNTQGENKFEVENLILHKDYSADTLAHNDIALLKIR-SYEG 258
 Db 393 NRPAPKELTVVLGQDRHNSQCEQCOTLAVRDYRLHFAFSPITYQH--DLALVRLQESADG 450
 QY 259 RCAQPSRTIQTICLSMYNDPOFGTS--CEITFGKENSSTLYPEQKMTWVKLISHRE 316
 Db 451 CCAHSPFPVQPCLESTAAAPAEAAVCEVAGWGHQFEGGE-YSSFLQEAQVPLIDPQR 509
 QY 317 COQPHYGSEVTTKMLCAADPOWKTDCQSGSGGPLVC---SLOGRWTLTGIVSWGRGCA 373
 Db 510 CSAPDVHGAFTQGMCLCAFLGEGTDACQSGSGGPLVCEDETPERQLILRGIVSWGSGCG 569
 QY 374 LKDKGVTVTRVSHFLPWIRSHST 395
 Db 570 NKLKGVTVTVANILAWIREHT 591

RESULT 21
 PLMN PIG
 ID PLMN PIG STANDARD; PRT; 790 AA.
 AC P06567;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasminogen (EC 3.4.21.7).
 GN PLG.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE OF 1-560.
 RA Schaller J., Marti T., Roessel S.J., Kaemper U., Rickli E.E.;
 RT "Amino acid sequence of the heavy chain of porcine plasmin. Comparison
 of the carbohydrate attachment sites with the human and bovine
 species".
 RL Fibrinolysis 1:91-102(1987).
 RN [2]
 RP SEQUENCE OF 450-790.
 RX MEDLINE=85203907; PubMed=3846533;
 RA Marti T., Schaller J., Rickli E.E.;
 RT "Determination of the complete amino-acid sequence of porcine
 miniplasminogen".
 RL Eur. J. Biochem. 149:279-285(1985).
 RN [3]
 RP CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=88185329; PubMed=3356193;
 RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
 RA Gerwig G.J., van Halbeek H., Vliegthart J.F.;
 RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
 plasminogen. Species specificity in relation to sialylation and
 fucosylation patterns".
 RL Eur. J. Biochem. 173:57-63(1988).
 CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
 A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
 EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
 AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
 GRAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
 AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
 LAMININ AND VON WILLEBRAND FACTOR.
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 higher selectivity than trypsin. Converts fibrin into soluble
 products.
 CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
 ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
 FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
 CC -!- PTM: N-LINKED GLYCAN CONTAINS N-ACETYLACTOSAMINE, SIALIC ACID AND
 IS CORE FUCOSYLATED. O-LINKED GLYCANS CONSIST OF GAL-GALNAc
 DISACCHARIDE WITH IS MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES
 (MICROHETEROGENEITY).
 CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
 IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -!- SIMILARITY: Contains 5 kringle domains.
 DR PIR; S03733; PLPG.
 DR HSP; P00747; SHPG.
 DR MEROPS; S01.233; .
 DR GlycoSuiteDB; P06867; .
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan.app.
 DR InterPro; IPR003966; Prothrombin.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00051; kringle; 5.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPsin.
 DR PRINTS; PR00018; KRINGLE.
 DR PRINTS; PR01505; PROTHROMBIN.
 DR ProDom; PD000395; Kringle; 5.
 DR SMART; SM00130; KR; 5.
 DR SMART; SM00473; PAN AP; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 5.
 DR PROSITE; PSS0070; KRINGLE_2; 5.
 DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat.
 KW CHAIN 1 560 PLASMIN HEAVY CHAIN A.
 FT CHAIN 561 790 PLASMIN LIGHT CHAIN B.
 FT DOMAIN 561 790 SERINE PROTEASE.
 FT KRINGLE 1.
 FT KRINGLE 2.
 FT KRINGLE 3.
 FT KRINGLE 4.
 FT KRINGLE 5.
 FT ACT_SITE 602 602 CHARGE RELAY SYSTEM.
 FT ACT_SITE 645 645 CHARGE RELAY SYSTEM.
 FT ACT_SITE 740 740 CHARGE RELAY SYSTEM.
 FT CARBOHYD 289 289 N-LINKED (GLCNAC...).
 FT CARBOHYD 340 340 /FTID=CAR_000019.
 FT CARBOHYD 340 340 O-LINKED (GALNAc...).
 FT /FTID=CAR_000020.
 SQ SEQUENCE 790 AA; 88592 MW; F04EA06E74BCD58E CRC64;
 Query Match 22.5%; Score 508.5; DB 1; Length 790;
 Best Local Similarity 36.2%; Pred. No. 6.7e-33;
 Matches 130; Conservative 45; Mismatches 141; Indels 43; Gaps 14;
 QY 45 DSKTCYEGNGHYRGKASDTWGRPCLPNSATVLOOTYHAHR-----SDALQLGLGKH 100
 Db 456 DLSEDCMFGNGKRYGRKATTVAGVPCQEWAA-----QEPHRSIFTPETNPRAGLEK-N 509
 QY 101 YCRNPD-NRRRRPWCVQVGLKPLVQECMVHDCADGKLKFCQGOKTLRPR---FKIIGEEF 156
 Db 510 YCRNPDGDDNGPCYT-TNQKLFIDYCDVQCYS--SFDCGKPKVKKPCARVVGCV 566
 QY 157 TTENQWPFAIYRRHRGGSVTVCGSLISPCWISATHCFIDYPKEDYIVYVLSRSL 216
 Db 567 SIPHSNWPQVSLRYRYRG----HFCGGLTISPWLVLTAKHCLKESSPSSPKVILGAHEE 622
 QY 217 NSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMY 276
 Db 623 YHLEGVQEDVSKLF--KEPS-----EADIALKLSSP-----AVITDKVIFACLP-- 668
 QY 277 NDPQF----QTSCEITGFGKENSTVDLYPEQLKMTVVKLISHRECQOPHYVGEVTTKML 332
 Db 669 --ENVVADVATACVIITGWGETKGT--YGAGLLKEARLPVIEKVCNRYEYLGKGVSPNEL 724
 QY 333 CAADPQWKTSCQDGGPLVCSLQGRMTLTGTVSGRCALXDKPGVTVRVSHFLPWI 391
 Db 725 CAGHLAGIDSCQDGGPLVCFEKKYILQGVTSWGLCALPNKPGVTVRVSRFTWI 783

RESULT 22
ID PLMN MOUSE STANDARD; PRT; 812 AA.
AC P2019; Q8CIS2; Q91WJ5;
DT 01-FEB-1991 (Rel. 17, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PLG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91194812; PubMed=2081600;
RA Degen S.J., Bell S.M., Schaefer L.A., Elliott R.W.;
RT "Characterization of the cDNA coding for mouse plasminogen and
RT localization of the gene to mouse chromosome 17.";
RL Genomics 8:49-61(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RA Brathwaite M., Waeltz P., Qian Y., Dudekula D., Schlessinger D.,
RA Nagaraja R.;
RT "Genomic sequence analysis in the mouse t-complex region.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 1-16 FROM N.A.
RC STRAIN=129/SVJ; TISSUE=Liver;
RX PubMed=12149246;
RA Bannach F.G., Gutierrez A., Fowler B.J., Bugge T.H., Degen J.L.,
RA Farmer R.J., Miles L.A.;
RT "Localization of regulatory elements mediating constitutive and
RT cytokine-stimulated plasminogen gene expression.";
RL J. Biol. Chem. 277:38579-38588(2002).
RN [5]
RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT suppression of metastases by a Lewis lung carcinoma.";
RL Cell 79:315-328(1994).
CC -1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,

CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC GRAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -1- FUNCTION: ANGIOSTATIN IS AN ANGIOGENESIS INHIBITOR THAT BLOCKS
CC NEOVASCULARIZATION AND GROWTH OF EXPERIMENTAL PRIMARY AND
CC METASTATIC TUMORS IN VIVO.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC -1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -1- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION
CC INVOLVES ONLY CLEAVAGE AFTER ARG-581, RESULTING IN 2 CHAINS HELD
CC TOGETHER BY 2 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE
CC ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -1- SIMILARITY: Contains 5 kringle domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J04766; AAA50168.1; -;
CC EMBL; AF481053; AAM22156.1; -;
CC EMBL; BC014773; AAI14773.1; -;
CC EMBL; AY134430; AAI15805.1; -;
CC FIR; A38514; PLMS.
CC HSP; P00747; 1PKX.
CC MEROPS; S01.233; -;
CC MGD; MGI:97620; Pig.
CC GO; GO:0016506; F:apoptosis activator activity; IDA.
CC GO; GO:0006915; P:apoptosis; IDA.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR003014; PAN.
CC InterPro; IPR003609; Pan_app.
CC InterPro; IPR003966; Prothrombin.
CC InterPro; IPR01254; Ser_protease_Try.
CC Pfam; PF00024; Kringle; 5.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC PRINTS; PR01505; PROTHROMBIN.
CC ProDom; PD000395; Kringle; 5.
CC SMART; SM00130; KR; 5.
CC SMART; SM00473; PAN_Ap; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00021; KRINGLE_1; 5.
CC PROSITE; PS00070; KRINGLE_2; 5.
CC PROSITE; PS0240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_His; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
KW Signal.
FT SIGNAL. 1 19
FT CHAIN 20 812 PLASMINOGEN.
FT CHAIN 20 581 PLASMIN HEAVY CHAIN A.
FT PEPTIDE 20 97 ACTIVATION PEPTIDE.
FT CHAIN 98 581 PLASMIN SHORT FORM OF CHAIN A.
FT CHAIN 98 2436 ANGIOSTATIN.
FT CHAIN 582 812 PLASMIN LIGHT CHAIN B.
FT DOMAIN 103 181 KRINGLE 1.

```
FT DOMAIN 184 262 KRINGLE 2.
FT DOMAIN 275 352 KRINGLE 3.
FT DOMAIN 377 454 KRINGLE 4.
FT DOMAIN 481 560 KRINGLE 5.
FT DOMAIN 582 812 SERINE PROTEASE.
FT ACT_SITE 624 624 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 667 667 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 762 762 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 49 73 BY SIMILARITY.
FT DISULFID 53 61 BY SIMILARITY.
FT DISULFID 103 181 BY SIMILARITY.
FT DISULFID 124 164 BY SIMILARITY.
FT DISULFID 152 176 BY SIMILARITY.
FT DISULFID 185 262 BY SIMILARITY.
FT DISULFID 206 245 BY SIMILARITY.
FT DISULFID 234 257 BY SIMILARITY.
FT DISULFID 275 352 BY SIMILARITY.
FT DISULFID 296 335 BY SIMILARITY.
FT DISULFID 324 347 BY SIMILARITY.
FT DISULFID 377 454 BY SIMILARITY.
FT DISULFID 398 437 BY SIMILARITY.
FT DISULFID 426 449 BY SIMILARITY.
FT DISULFID 481 560 BY SIMILARITY.
FT DISULFID 502 543 BY SIMILARITY.
FT DISULFID 531 555 BY SIMILARITY.
FT DISULFID 568 687 INTERCHAIN (BY SIMILARITY).
FT DISULFID 578 586 INTERCHAIN (BY SIMILARITY).
FT DISULFID 609 625 BY SIMILARITY.
FT DISULFID 701 768 BY SIMILARITY.
FT DISULFID 731 747 BY SIMILARITY.
FT DISULFID 758 786 BY SIMILARITY.
FT CONFLICT 235 235 R -> H (IN REF. 1).
FT CONFLICT 525 525 S -> D (IN REF. 1).
FT CONFLICT 649 649 G -> L (IN REF. 1).
SQ SEQUENCE 812 AA; 90781 MW; 2417326056A2FFD2 CRC64;

Query Match 22.4%; Score 505; DB 1; Length 812;
Best Local Similarity 33.2%; Pred. No. 1.3e-32;
Matches 135; Conservative 48; Mismatches 170; Indels 54; Gaps 15;

QY 10..NCDCLNGTGVSNKYFSNIHWCNCPK--KFGQHCIEI-----DKSKTCYEGNGH 56
DB 428 NPDGDKGWCYTTDPSVWEVCNLCRCSETGSGVVELPTVQEPGSDSETDCMYGNXK 487
QY 57 FYRGKASTDTWGRPLCPNWSATVLOQTVHARSALQ-----LGIGKKNYCNRPD-NRRRP 111
DB 488 DYRGKTAVTAAAGTFCQGWAA-----QEPHRSIETPTQTNPRAGLEK-NYCRNPDGVNRP 541
QY 112 WCYVQVGLKPLVQECMVHDCADGKLKFCQGGOKTLRPR---FKIIGGEFTTIENOPWFAAI 168
DB 542 WCYT-TNPRKLYDYCDIPLCASAS-SFECGRQVPEPKCPGRVVGCVANPHSPWQISL 599
QY 169 YRRHGGSVTVVCGSLISPCWVLSATHCFIDYPKKEDYIYVLRGSRNLNSNTQGMKFEV 228
DB 600 RTRFTG---QHFCGGTLIAPEVLTAAHCLKESRPEPKYKVLGAHBEYIRGSDVQEISV 656
QY 229 ENLLHKDYSADTLAHNDIALKIRSKEGCAQPSRTIQICLPSMNDPQF----GTS 284
DB 657 AKLLE-----PNNRDIALKLU-----SPAITDKVIFACLPSS-----PNYVWADRTI 700
QY 285 CEITGFGKENGSTDYLPBQLKXVTVKLISHRECOQPHYVGGSEVTTKMLCAADPQWKTDSC 344
DB 701 CVITGWGETQGT--FGAGRLKEAQLPVLENKCNREYLNKRVKSTELCAQLAGGVGVDSC 758
QY 345 QGDSGGPLVCSLQGMTLITGVSGRCALKDKCGVYTRVSHFLPWI 391
DB 759 QGDSGGPLVCFEXDKYIYLGQVTSWGLGCARENKPKGVYVVRVSRFDWI 805

RESULT 23
PLMN_MACMU
ID PLMN_MACMU STANDARD; PRG: 810 AA.
```

PROSITE; PS00135; TRYPSIN SER; 1.
Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
Signal.

FT	SIGNAL	1	19	PLASMINOGEN.
FT	CHAIN	20	810	PLASMIN HEAVY CHAIN A.
FT	CHAIN	20	580	ACTIVATION PEPTIDE.
FT	PEPTIDE	20	96	PLASMIN SHORT FORM OF CHAIN A.
FT	CHAIN	97	580	PLASMIN LIGHT CHAIN B.
FT	CHAIN	581	810	KRINGLE 1.
FT	DOMAIN	103	181	KRINGLE 2.
FT	DOMAIN	184	262	KRINGLE 3.
FT	DOMAIN	275	352	KRINGLE 4.
FT	DOMAIN	377	454	KRINGLE 5.
FT	DOMAIN	481	560	SERINE PROTEASE.
FT	DOMAIN	581	810	CHARGE RELAY SYSTEM.
FT	ACT SITE	632	632	CHARGE RELAY SYSTEM.
FT	ACT SITE	665	665	CHARGE RELAY SYSTEM.
FT	ACT SITE	760	760	CHARGE RELAY SYSTEM.
FT	BINDING	136	136	OMEGA-AMINOCARBOXYLIC ACIDS.
FT	BINDING	158	158	OMEGA-AMINOCARBOXYLIC ACIDS.
FT	BINDING	172	172	OMEGA-AMINOCARBOXYLIC ACIDS.
FT	BINDING	432	432	OMEGA-AMINOCARBOXYLIC ACIDS.
FT	BINDING	445	445	OMEGA-AMINOCARBOXYLIC ACIDS.
FT	BINDING	134	134	OMEGA-AMINOCARBOXYLIC ACIDS.
FT	BINDING	136	136	FIBRIN.
FT	DISULFID	49	73	BY SIMILARITY.
FT	DISULFID	53	61	BY SIMILARITY.
FT	DISULFID	103	181	BY SIMILARITY.
FT	DISULFID	124	164	BY SIMILARITY.
FT	DISULFID	152	176	BY SIMILARITY.
FT	DISULFID	185	262	BY SIMILARITY.
FT	DISULFID	188	316	BY SIMILARITY.
FT	DISULFID	205	245	BY SIMILARITY.
FT	DISULFID	234	257	BY SIMILARITY.
FT	DISULFID	275	352	BY SIMILARITY.
FT	DISULFID	296	335	BY SIMILARITY.
FT	DISULFID	324	347	BY SIMILARITY.
FT	DISULFID	377	454	BY SIMILARITY.
FT	DISULFID	398	437	BY SIMILARITY.
FT	DISULFID	426	449	BY SIMILARITY.
FT	DISULFID	481	560	BY SIMILARITY.
FT	DISULFID	502	543	BY SIMILARITY.
FT	DISULFID	531	555	BY SIMILARITY.
FT	DISULFID	567	685	BY SIMILARITY.
FT	DISULFID	577	585	BY SIMILARITY.
FT	DISULFID	607	623	BY SIMILARITY.
FT	DISULFID	699	766	BY SIMILARITY.
FT	DISULFID	729	745	BY SIMILARITY.
FT	DISULFID	756	784	BY SIMILARITY.
FT	CARBOHYD	365	365	BY SIMILARITY.
FT	SEQUENCE	810	AA:	O- ¹⁵ SEIC5A1A0P24A CRC64:
FT	SEQUENCE	810	AA:	A- ¹⁵ SEIC5A1A0P24A CRC64:
FT	SEQUENCE	810	AA:	BY SIMILARITY.

```

Qy 229 ENLIHKOYSADTLAHHNDIALLKIRSKRGCAQPSRTIOTICLPWSYNDPQF-----GTS 288
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 650 QBEIVSKMFSEPARA---DIALKLSSP---AIITDKVIPACLLPS-----PNTVVADRT 698
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 285 CEITGFGKENSFDYILPSPQLKMTVKVLISHRECOQPHYGVSEVTKVLCAADPQWKTDSC 344
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 699 CFIITGWETQGT--YGAGLLKEALPVIENKVCNRYEFLNGTVKTTLCAGHLAGGTDSC 756
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 345 QGDSGGPLVCSLQGRMTLTIGVSGRCALKDKKPGVTVRVSHFLPWI 391
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 757 QGDSGGPLVCEPKDKYILQGVTSNGLGCAEPKPKGVVVRVSRFTWI 803
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 24
PLMN HUMAN
ID PLMN HUMAN STANDARD; PRT; 810 AA.
AC P00747;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PUG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=90202879; PubMed=2318948;
RA Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
   in the fibrinolytic system.";
RL J. Biol. Chem. 265:6104-6111(1990).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=87162490; PubMed=3030813;
RA Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
   for human plasminogen.";
RL FEBS Lett. 213:254-260(1987).
[3]
RN SEQUENCE OF 20-810.
RA Sottrup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the EMBL data bank.
[4]
RN SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Walinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
   human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
[5]
RN SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
   of human plasminogen and their interaction with the NH2-terminal
   activation peptide as studied by affinity chromatography.";
RL Eur. J. Biochem. 50:489-494(1975).
[6]
RN SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810.
RA Sottrup-Jensen L., Claeys H., Zajdel M., Petersen T.E., Magnusson S.;
RL (in) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);
RL Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,
   Raven Press, New York (1978).
[7]
RN SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126863;
RA Wiman B., Wallen P.;
RT "Amino-acid sequence of the cyanogen-bromide-fragment from human
   plasminogen that forms the linkage between the plasmin chains.";
RL Eur. J. Biochem. 58:539-547(1975).
[8]

```

RP SEQUENCE OF 581-810.
RX MEDLINE=77225245; PubMed=142009;
RA Wiman B.;
RT "Primary structure of the B-chain of human plasmin.";
RL Eur. J. Biochem. 76:129-137(1977).
RN [9]
RP ACTIVE SITE
RX MEDLINE=73149248; PubMed=4894729;
RA Robbins K.C., Bernabe P., Arzadon L., Summaria L.;
RT "The primary structure of human plasminogen. II. The histidine loop
of human plasmin: light (B) chain active center histidine sequence.";
RL J. Biol. Chem. 248:1631-1633(1973).
RN [10]
RP ACTIVE SITE
RX MEDLINE=69234739; PubMed=4240117;
RA Grobkopf W.R., Summaria L., Robbins K.C.;
RT "Studies on the active center of human plasmin. Partial amino acid
sequence of a peptide containing the active center serine residue.";
RL J. Biol. Chem. 244:3590-3597(1969).
RN [11]
RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
RX MEDLINE=82213905; PubMed=6919539;
RA Trexler M., Vali Z., Parthy L.;
RT "Structure of the omega-aminocarboxylic acid-binding sites of human
plasminogen. Arginine 70 and aspartic acid 56 are essential for
binding of ligand by kringle 4.";
RL J. Biol. Chem. 257:7401-7406(1982).
RN [12]
RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
RX MEDLINE=85054794; PubMed=6094526;
RA Valli Z., Parthy L.;
RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34
are essential for fibrin affinity of the kringle 1 domain.";
RL J. Biol. Chem. 259:13690-13694(1984).
RN [13]
RP PHOSPHORYLATION SITE SER-597.
RX MEDLINE=97345939; PubMed=9201958;
RA Wang H., Prorok M., Brethauer R.K., Castellino F.J.;
RT "Serine-578 is a major phosphorylation locus in human plasma
plasminogen.";
RL Biochemistry 36:8100-8106(1997).
RN [14]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=88185329; PubMed=3356193;
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
plasminogen. Species specificity in relation to sialylation and
fucosylation patterns.";
RL Eur. J. Biochem. 173:57-63(1988).
RN [15]
RP CARBOHYDRATE-LINKAGE SITE SER-268.
RX MEDLINE=97207306; PubMed=9054441;
RA Pirie-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,
Pizzo S.V.;
RT "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of
human plasminogen 2.";
RL J. Biol. Chem. 272:7408-7411(1997).
RN [16]
RP CHARACTERIZATION OF ANGIOSTATIN AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shang Y., Chen C., Rosenthal R.A.,
Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
suppression of metastases by a Lewis lung carcinoma.";
RL Cell 79:315-328(1994).
RN [17]
RP CHARACTERIZATION OF ANGIOSTATIN.
RX MEDLINE=97238710; PubMed=9102221;
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madesen J.W.,
Lapcevich R., Nacy C.A.;
RT "A recombinant human angiotensin protein inhibits experimental primary
and metastatic cancer.";

RL Cancer Res. 57:1329-1334(1997).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
RX MEDLINE=92031502; PubMed=1657148;
RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;
RT "Crystal and molecular structure of human plasminogen kringle 4
refined at 1.9-A resolution.";
RL Biochemistry 30:10576-10588(1991).
RN [19]
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
RX MEDLINE=92031503; PubMed=1657149;
RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
RT "The refined structure of the epsilon-aminocaproic acid complex of
human plasminogen kringle 4.";
RL Biochemistry 30:10589-10594(1991).
RN [20]
RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
RA Stec B., Yamano A., Whitlow M., Teeter M.W.;
RT "Structure of human plasminogen kringle 4 at 1.68 Angstrom and 277 K.
A possible structural role of disordered residues.";
RL Acta Crystallogr. D 53:169-178(1997).
RN [21]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.
RX MEDLINE=96180681; PubMed=8611560;
RA Mathews I.I., Vanderhoff-Hanaver P., Castellino F.J., Tulinsky A.;
RT "Crystal structures of the recombinant kringle 1 domain of human
plasminogen in complexes with the ligands epsilon-aminocaproic acid
and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
RL Biochemistry 35:2567-2576(1996).
RN [22]
RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
RX MEDLINE=98198034; PubMed=9521645;
RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
Castellino F.J.;
RT "Structure and ligand binding determinants of the recombinant kringle
5 domain of human plasminogen.";
RL Biochemistry 37:3258-3271(1998).
RN [23]
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237157; PubMed=8181475;
RA Rejante M.R., Llinas M.;
RT "1H-NMR assignments and secondary structure of human plasminogen
kringle 1.";
RL Eur. J. Biochem. 221:927-937(1994).
RN [24]
RP STRUCTURE BY NMR OF 96-184.
RX MEDLINE=94237158; PubMed=8181476;
RA Rejante M.R., Llinas M.;
RT "Solution structure of the epsilon-aminohexanoic acid complex of
human plasminogen kringle 1.";
RL Eur. J. Biochem. 221:939-949(1994).
RN [25]
RP STRUCTURE BY NMR OF 183-354.
RX MEDLINE=96194156; PubMed=8652577;
RA Soehnle S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
Rickli E.E.;
RT "Recombinant gene expression and 1H NMR characteristics of the
kringle (2 + 3) supermodule: spectroscopic/functional individuality
of plasminogen kringle domains.";
RL Biochemistry 35:2357-2364(1996).
RN [26]
RP STRUCTURE BY NMR OF 374-461.
RX MEDLINE=90219023; PubMed=2157850;
RA Atkinson R.A., Williams R.J.P.;
RT "Solution structure of the kringle 4 domain from human plasminogen by
1H nuclear magnetic resonance spectroscopy and distance geometry.";
RL J. Mol. Biol. 212:541-552(1990).
RN [27]
RP VARIANTS PHE-374 AND THR-620.
Query Match 22.1%; Score 499; DB 1; Length 810;
Best Local Similarity 33.7%; Pred. No. 3,9e-32;
Matches 137; Conservative 44; Mismatches 170; Indels 56; Gaps 15;

QY 10 NCDCLNGGTCUSNKFYSNIHWNCNPKKFGQ-----HCEIDSKTCYBGNH 56
 Db 428 NPDADKGPWCFTTDPVSRWYCNLKKSCGTEASVAPPVLLPDVETPSEEDCMFGNGK 497
 QY 57 FYRGKAGTDMGRPCLPWNATVLQOYTHAHR-----SDALQLGLGKHNCRNPD-NRRRP 111
 Db 488 GYRGKRAATTGTGFCQDMAA-----QEPHRSIFTPTNPQAGLEK-NYCRNPDGCVGGP 541
 QY 112 WCYVQVGLKPLVQSCVHDCADGKLFQCCQKTLRPR---FKIIGBFTTIENQPFPAI 168
 Db 542 WCYT-TNPRKLYDYCDVFEQCA--APSEDCGKQPVEPKKCPGRVVGCVAAHPHSPWQVSL 598
 QY 169 YRRHRGGSVTVCGSLSPCWISATHCFIDYPKKEDYVYLCRSRLNNTQCEMKFEV 228
 Db 599 --RTRFG--MHFCGGTLLISEWVLTAAHCLKSPSSYKVLGAHCEVNLPEHVQIEV 654
 QY 229 ENLILHKDYSADTLAHHNDIALKIRSKRGCAQPSRTIOTICLPSWYNDPQF-----GTS 284
 Db 655 SRLFLEPT-----RKDIALLLKSSP---AVITDKVIPACLPS-----PNYVADRT 698
 QY 285 CEITGFGKENSTDYLPQELKMTVVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSC 344
 Db 699 CFITGGETQGT--FGAGLLKEAQLPVNIENKVCNRYEFLNGRVOSTELCAGHLAGGTDSC 756
 QY 345 QGDSGGPLVCSLQGRMTLTGIVSGRGCAKDKPGVTVRVSHFLPWI 391
 Db 757 QGDSGGPLVCFEKDKYILQGVTSWGLGCARENKPGVTVRVSRFTWI 803

RESULT 25

PLMN CANFA STANDARD; PRT; 333 AA.
 ID - PLMN CANFA
 AC P80009;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Plasminogen (EC 3.4.21.17) (Fragment).
 GN PLG.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_taxID=9615;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Plasma;
 RX MEDLINE=90175323; PubMed=2626424;
 RA Schaller J., Straub C., Kaempfer U., Rickli E.E.;
 RT "Complete amino acid sequence of canine miniplasminogen.";
 RL Protein Seq. Data Anal. 2:445-450(1989).
 CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
 CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
 CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
 CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
 CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS. SUCH
 CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
 CC LAMININ AND VON WILLEBRAND FACTOR.
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 CC higher selectivity than trypsin. Converts fibrin into soluble
 CC products.
 CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
 CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
 CC FIBRIN. ACTIVATED WITH UROKINASE AND HIGH CONCENTRATIONS OF
 CC STREPTOKINASE.
 CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
 CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -!- SIMILARITY: Contains at least 1 kringle domain.
 CC HSSP; P00747; SHPG.
 DR MEROPS; S01.233; .
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003966; Prothrombin.
 DR InterPro; IPR001254; Ser. protease_Try.
 DR Pfam; PF00051; kringle_1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PRO0722; CHYMOTRYPSIN.
 DR PRINTS; PRO0018; KRINGLE.
 DR PRINTS; PRO1505; PROTHROMBIN.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00130; K3; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 KW Tissue remodeling; Blood coagulation; Kringle; Zymogen.
 FT NON_TER 1
 FT CHAIN <1 103 PLASMIN HEAVY CHAIN A.
 FT CHAIN 104 333 PLASMIN LIGHT CHAIN B.
 FT DOMAIN 4 83 KRINGLE 5.
 FT DOMAIN 104 333 SERINE PROTEASE.
 FT DISULFID 4 83 BY SIMILARITY.
 FT DISULFID 25 66 BY SIMILARITY.
 FT DISULFID 54 78 BY SIMILARITY.
 FT DISULFID 90 208 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 100 108 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 130 146 BY SIMILARITY.
 FT DISULFID 222 289 BY SIMILARITY.
 FT DISULFID 252 268 BY SIMILARITY.
 FT DISULFID 279 307 BY SIMILARITY.
 FT ACT_SITE 145 145 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 188 188 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 283 283 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT SITE 152 152 STREPTOKINASE-BINDING SITE (PROBABLE).
 FT SITE 186 186 STREPTOKINASE-BINDING SITE (PROBABLE).
 FT SITE 264 264 STREPTOKINASE-BINDING SITE (PROBABLE).
 FT SITE 277 277 SITE OF SUBSTRATE SPECIFICITY (BY SIMILARITY).
 SQ SEQUENCE 333 AA; 36678 MW; C8C0271B6C6AC8D4 CRC64;
 Query Match 22.1%; Score 498.5; DB 1; Length 333;
 Best Local Similarity 35.6%; Pred. No. 1-6e-32;
 Matches 126; Conservative 49; Mismatches 136; Indels 43; Gaps 14;
 QY 50 CYEGNGHFYRGKASTDTMGRPCLPWNATVLQOYTHAHR-----SDALQLGLGKHNCRNP 105
 Db 4 CMFNGKGYRGKATTVMGIPQCEWAA-----QEPHRSIFTPTNPQAGLEK-NYCRNP 57
 QY 106 D-NRRRPPWCYVQVGLKPLVQSCVHDCADGKLFQCCQKTLRPR---FKIIGBFTTIEN 161
 Db 58 DGVNGPWCYT-MNQKLFYDCVQCV--STSFDCGKQVPEPKKCPGRVVGCVANPHS 114
 QY 162 QPMFAAIYRRHRGGSVTVCGSLSPCWISATHCFIDYPKKEDYVYLCRSRLNNTQ 221
 Db 115 WPMQISLRTRYG---KHFCGGTLLISEWVLTAAHCLERSRRSPASYKVLG-----AHKE 165
 QY 222 GEMKPFVENLILHKDYSADTLAHHNDIALKIRSKRGCAQPSRTIOTICLPSWYNDPQF 281
 Db 166 VNLESQVQIEVYKLFLEPTRA---DIALLLKSSP---AVITSKVIPACLPS-----PPNY 214
 QY 282 ----GTSCEITGFGKENSTDYLPQELKMTVVKLISHRECQPHYGVSEVTTKMLCAADP 337
 Db 215 VVADRTLCTVITGWGETQGT--YGAGLLKEAQLPVNIENKVCNRYEFLNGRVOSTELCAGNL 272
 QY 338 QWKTDSCQDSGGPLVCSLQGRMTLTGIVSGRGCAKDKPGVTVRVSHFLPWI 391
 Db 273 AGGTDSCQDSGGPLVCFEKDKYILQGVTSWGLGCARENKPGVTVRVSRFTWI 326

RESULT 26

PLMN SHEEP
 ID - PLMN_SHEEP STANDARD; PRT; 343 AA.

AC	P81286;	50	CYEGNGHFYRGKASTDTMGRPCLPNSATVLQOQTYHAH----	RSDALQLGLCKHNYCRNP	105
DT	15-DEC-1998 (Rel. 37, Created)	15	CMLGKGYRGKATVAGVPCQEWAA-----QEPHRRGIFTPEINPRAGLEK-NYCRNP	68	
DT	15-DEC-1998 (Rel. 37, Last sequence update)	106	D-NRRRPMCYVQVGLKPLVQECMHDCADGKLFQCGOKTLRPR---FKLIGGETTIEEN	161	
DE	28-FEB-2003 (Rel. 41, Last annotation update)	69	DGDVANGPCYT-TNPRKLFYCDIPQC---BSSFDGPKPVEPKKPCPARVVGCVATPHS	124	
GN	PLG.	162	QWFAAIYRRHGGSVTVYCGSLSPQWVISATHCFIDYPKKEDYIVVIGRSRLNSNQ	221	
OS	Ovis aries (Sheep).	125	WPQVSLRRRSR-----EHFCGGTLLSPWLVLTAAHCLDSILGSPSYTVILGAHYEMAREA	180	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	222	GEMKFEVENLILHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPQF	281	
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;	181	SVQEPVSELFLEPSRA-----DIALLLSSP-----AVITDEVIPACLP	224	
OC	Bovidae; Caprinae; Ovis.	282	-----GTSCHITGFKENSTDYLPQLKMTVVKLISHRECOQPHYGVSEVITKMLCAADP	337	
NCBI	TaxID=9940;	225	VVADKTVCYITGWTGTGQF--FVGRLKEARLPVIEKNVYCNRYELNGRVKSTELCAGDL	282	
FN	SEQUENCE	338	QWKTSCQDGGPLVCSLQGRMTLTGVSMRGCAKDKPGVYTRVSHFLPMI	391	
RA	MEDLINE=93149995; PubMed=1492092;	283	AGTDSQDGGPLVCFEKDYILQGVTSWGLGCARPKNKPGVYRVSTYVPMI	336	
RA	Schaller J., Straub C., Kampfer U., Rickli E.E.;				
RT	"Complete amino acid sequence of ovine miniplasminogen.";				
RL	Protein Seq. Data Anal. 5:21-25(1992).				
CC	-!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS				
CC	A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING				
CC	EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,				
CC	AND INFLAMMATION. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN				
CC	GRAFTIAN FOLLICLE. IT OVULATION IT WEAKENS THE WALLS OF THE				
CC	ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH				
CC	AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,				
CC	LAMININ AND VON WILLEBRAND FACTOR.				
CC	-!- CATALYTIC ACTIVITY: Preferential cleavage: Lys- -Xaa > Arg- -Xaa;				
CC	higher selectivity than trypsin. Converts fibrin into soluble				
CC	products.				
CC	-!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN				
CC	ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO				
CC	FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.				
CC	-!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN				
CC	IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.				
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.				
CC	-!- SIMILARITY: Contains at least 2 kringle domains.				
DR	PIR, B61545; B61545.				
DR	HSPF, P00747; SHPG.				
DR	MEROPS; S01.233; --				
DR	InterPro; IPR001314; Chymotrypsin.				
DR	InterPro; IPR000001; Kringle.				
DR	InterPro; IPR003966; Prothrombin.				
DR	InterPro; IPR001254; Ser. protease_Try.				
DR	Pfam; PF00051; Kringle; 1.				
DR	Pfam; PF00089; trypsin; 1.				
DR	PRINTS; PR00722; CHYMOTRYPSIN.				
DR	PRINTS; PR00018; KRINGLE.				
DR	PRINTS; PR01505; PROTHROMBIN.				
DR	ProDom; PD000395; Kringle; 1.				
DR	SMART; SM00130; KR; 1.				
DR	SMART; SM00020; Tryp SPC; 1.				
DR	PROSITE; PS00021; KRINGLE_1; 1.				
DR	PROSITE; PS00070; KRINGLE_2; 1.				
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.				
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.				
DR	PROSITE; PS00135; TRYPSIN_SER; 1.				
KW	Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;				
KW	Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat.				
FT	NON_TER	1			
FT	DOMAIN	<1	140	HEAVY CHAIN A.	
FT	DOMAIN	141	>343	LIGHT CHAIN A.	
FT	DOMAIN	<1	17	KRINGLE 4.	
FT	DOMAIN	41	120	KRINGLE 5.	
FT	DOMAIN	114	341	SERINE_PROTEASE.	
FT	ACT_SITE	181	181	CHARGE RELAY SYSTEM.	
FT	ACT_SITE	224	224	CHARGE RELAY SYSTEM.	
FT	ACT_SITE	319	319	CHARGE RELAY SYSTEM.	
FT	NON_TER	343	343		
SQ	SEQUENCE	343	AA; 37662 MW; 8DF6EBA92D596EE0	CRC64;	
Query Match 21.8%; Score 493; DB 1; Length 343;					
Best Local Similarity 35.6%; Pred. No. 4.5e-32;					
Matches 126; Conservative 42; Mismatches 142; Indels 44; Gaps 13;					

CC GRAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC AS C1 AND C3. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS. BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC FIBRIN: N-LINKED GLYCAN CONTAIN N-ACETYLGLUCOSAMINE AND SIALIC ACID.
CC O-LINKED GLYCANS CONSIST OF GAL-GALNAc DISACCHARIDE WITH IS
CC MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES (MICROHETEROGENEITY).
CC MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC SIMILARITY: Contains 5 kringle domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X79402; CAA5939.1; --
CC EMBL; K02935; AAA30714.1; --
CC PIR; S45046; PLK.
CC HSP; P00747; 2PK4.
CC MEROPS; S01.233; P06869; --
CC GlycoSuiteDB; P06869; --
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR003014; PAN.
CC InterPro; IPR003609; Pan app.
CC InterPro; IPR003966; Prothrombin.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00051; Kringle; 5.
CC Pfam; PF00024; PAN; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PF00722; CHYMOTRYPSIN.
CC PRINTS; PR00108; KRINGLE.
CC PRINTS; PR01505; PROTHROMBIN.
CC ProDom; PD000395; Kringle; 5.
CC SMART; SM00130; KR; 5.
CC SMART; SM00473; PAN_AP; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00021; KRINGLE_1; 5.
CC PROSITE; PS00070; KRINGLE_2; 5.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
CC Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
CC Signal.
FT SIGNAL 1 26
FT CHAIN 27 812 PLASMINOGEN
FT CHAIN 27 583 PLASMIN HEAVY CHAIN A.
FT CHAIN 27 583 PLASMIN LIGHT CHAIN B.
FT DOMAIN 584 812 KRINGLE 1.
FT DOMAIN 110 188 KRINGLE 2.
FT DOMAIN 192 269 KRINGLE 3.
FT DOMAIN 282 359 KRINGLE 4.
FT DOMAIN 384 461 KRINGLE 5.
FT DOMAIN 485 564 KRINGLE 5.
FT DOMAIN 584 812 SERINE PROTEASE.
FT CARBOHYD N-LINKED (GLCNAC. . .).
FT CARBOHYD /FTid=CAR_000014.
FT CARBOHYD O-LINKED (GALNAc. . .).
FT CARBOHYD /FTid=CAR_000015.
FT 365
FT ACT_SITE 624
FT ACT_SITE 667 CHARGE RELAY SYSTEM.
FT ACT_SITE 762 CHARGE RELAY SYSTEM.

FT CONFLICT 335 335 N -> D (IN REF. 2).
FT CONFLICT 516 516 Q -> H (IN REF. 2).
FT CONFLICT 555 555 P -> L (IN REF. 2).
FT CONFLICT 744 744 T -> R (IN REF. 3).
SQ SEQUENCE 812 AA; 91216 MW; 386AA691E220946 CRC64;
Query Match 21.5%; Score 484.5; DB 1; Length 812;
Best Local Similarity 35.4%; Pred. No. 5.6e-31;
Matches 123; Conservative 42; Mismatches 151; Indels 31; Gaps 11;
QY 50 CYEGNGHFVRGKASTDTMGRCPLPNVSATVLTQOTVHAHRSDALQGLGKHVCRNPD-NR 108
DB 485 CMIGTKSVRGKATTVAGVPCQEWAEQPHSHLFTETNP-QSGL-ERNVCRNPDDGV 542
QY 109 RRPWCYVQVGLKPLVQECMHDCADGKLFQCGQKTLRPR---FKIIGGEFTTIENQPMF 165
DB 543 NGPWCYTNPRPF-DYCDVPC---ESSFDCGKPKVEPKCGRIVGCGVSKPHSWPMQ 598
QY 166 AAIYRRHGGSVYVCGSLSPCVISATHCFIDYPKKEDIVVGLGRSLNSNTGEMK 225
DB 599 VSLRRSR-----HFCGGTLSPKWLTAHCLDNILALSPFYKILGAHNEKVRQSVQE 653
QY 226 FEVENLILHKDYSADTLAHNDIALIKIRSKRGCAQPSRTIQTICLPSMYNDPQGTSC 285
DB 654 IYVSLFRPESA-----DIALKL-----SRPAITKEVIPACLPNPNYVAARTEC 701
QY 286 EITGFGKENSIDYLYPE-OLKMTVVKLISHRCQPHYGVSEVTKMLCAADPQWKTDSC 344
DB 702 YITGWGETQGT---FQEGLLKEAHLFVIEKCNREYLDGRVKXPTELCAGHLIGTQDSC 758
QY 345 QGDSGGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPMI 391
DB 759 QGDSGGPLVCFEKKYKILQSVTISWGLGCARPNKPGYVYRVSPYVPMI 805
RESULT 28
APOA HUMAN
ID APOA_HUMAN STANDARD; PRT; 4548 AA.
AC P08519; 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Apolipoprotein(a) precursor (EC 3.4.21.-) (Apo(a)) (Lp(a)).
GN LPA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88039109; PubMed=3670400;
RA McLean J.W., Tomlison J.E., Kuang W.-J., Eaton D.L., Chen E.Y.,
RA Fless G.M., Scanu A.M., Lawn R.M.;
RT "cDNA sequence of human apolipoprotein(a) is homologous to
RT plasminogen.";
RL Nature 330:132-137(1987).
RN [2]
RP SERINE PROTEASE ACTIVITY.
RX MEDLINE=90076123; PubMed=2531657;
RA Saionen E.-M., Jaunainen M., Zardi L., Vaheri A., Ehnholm C.;
RT "Lipoprotein(a) binds to fibronectin and has serine proteinase
RT activity capable of cleaving it.";
RL EMBO J. 8:4035-4040(1989).
RN [3]
RP REVIEW.
RX MEDLINE=90049223; PubMed=2530631;
RA Utermann G.;
RT "The mysteries of lipoprotein(a)."
RL Science 246:904-910(1989).
RN [4]
RP CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.
RX MEDLINE=21303595; PubMed=11294842;
RA Garner B., Merry A.H., Royle L., Harvey D.J., Rudd P.M., Thillet J.;

RT "Structural elucidation of the N- and O-glycans of human
RT apolipoprotein(a): role of o-glycans in conferring protease
RT resistance."; J. Biol. Chem. 276:22200-22208(2001).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.
RX MEDLINE=96217891; PubMed=8642595;
RA Mikol V., Lograsso P.V., Boettcher B.R.;
RT "Crystal structures of apolipoprotein(a) kringle IV37 free and
RT complexed with 6-aminocaproic acid and with p-aminomethylbenzoic
RT acid: existence of novel and expected binding modes."; J.
RL J. Mol. Biol. 256:751-761(1996).
RN [6]
RP VARIANT ARG-4193.
RX MEDLINE=95002201; PubMed=7918682;
RA Scanu A.M., Pfaffinger D., Lee J.C., Hinman J.;
RT "A single point mutation (Trp2-->Arg) in human apo(a) kringle 4-37
RT associated with a lysine binding defect in Lp(a).";
RL Biochim. Biophys. Acta 1227:41-45(1994).
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp 330.
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC decorin.
CC -!- PTM: N- and O-glycosylated. The N-glycans are complex biantennary
CC structures present in either a mono- or disialylated state. The
CC O-glycans are mostly (80%) represented by the monosialylated core
CC type I structure, NeuNAcaphas-3Galbeta1-3GalNAc, with smaller
CC amounts of disialylated and non-sialylated O-glycans also
CC detected.
CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its
CC naturally occurring proteolytic fragments are correlated with
CC atherosclerosis. Homology with plasminogen kringles IV and V is
CC thought to underlie the atherogenicity of the protein, because the
CC fragments are competing with plasminogen for fibrin(ogen) binding.
CC -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
CC leading to the formation of the so called mini-Lp(a). Apo(a)
CC fragments accumulate in atherosclerotic lesions, where they may
CC promote thrombogenesis. O-glycosylation may limit the extent of
CC proteolytic fragmentation.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -!- SIMILARITY: Contains 38 kringle domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X06290; CA229618.1; -.
DR PIR; S00657; S00657.
DR PDB; 1I71; 13-JUN-01.
DR PDB; 1JFN; 28-JUN-02.
DR PDB; 1KIV; 18-MAY-99.
DR PDB; 3KIV; 18-MAY-99.
DR PDB; 4KIV; 18-MAY-99.
DR MEROPS; S01.226; -.
DR Genew; HGNC:6667; LPA.
DR MIM; 152200; -.
DR GO; GO:0004866; F:endopeptidase inhibitor activity; TAS.
DR GO; GO:0008015; P:circulation; TAS.
DR GO; GO:0009405; P:pathogenesis; TAS.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Serine protease_Try.
DR Pfam; PF00051; kringle; 38.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 38.

DR SMART; SM00130; KR; 38.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 38.
DR PROSITE; PS00070; KRINGLE_2; 38.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
KW Kringle; Repeat; Atherosclerosis; Signal; Polymorphism; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 4548 APOLIPOPROTEIN(A)
FT DOMAIN 20 130 KRINGLE TYPE IV, 1.
FT DOMAIN 131 244 KRINGLE TYPE IV, 2.
FT DOMAIN 245 358 KRINGLE TYPE IV, 3.
FT DOMAIN 359 472 KRINGLE TYPE IV, 4.
FT DOMAIN 473 586 KRINGLE TYPE IV, 5.
FT DOMAIN 587 700 KRINGLE TYPE IV, 6.
FT DOMAIN 701 814 KRINGLE TYPE IV, 7.
FT DOMAIN 815 928 KRINGLE TYPE IV, 8.
FT DOMAIN 929 1042 KRINGLE TYPE IV, 9.
FT DOMAIN 1043 1156 KRINGLE TYPE IV, 10.
FT DOMAIN 1157 1270 KRINGLE TYPE IV, 11.
FT DOMAIN 1271 1384 KRINGLE TYPE IV, 12.
FT DOMAIN 1385 1498 KRINGLE TYPE IV, 13.
FT DOMAIN 1499 1612 KRINGLE TYPE IV, 14.
FT DOMAIN 1613 1726 KRINGLE TYPE IV, 15.
FT DOMAIN 1727 1840 KRINGLE TYPE IV, 16.
FT DOMAIN 1841 1954 KRINGLE TYPE IV, 17.
FT DOMAIN 1955 2068 KRINGLE TYPE IV, 18.
FT DOMAIN 2069 2182 KRINGLE TYPE IV, 19.
FT DOMAIN 2183 2296 KRINGLE TYPE IV, 20.
FT DOMAIN 2297 2410 KRINGLE TYPE IV, 21.
FT DOMAIN 2411 2524 KRINGLE TYPE IV, 22.
FT DOMAIN 2525 2638 KRINGLE TYPE IV, 23.
FT DOMAIN 2639 2752 KRINGLE TYPE IV, 24.
FT DOMAIN 2753 2866 KRINGLE TYPE IV, 25.
FT DOMAIN 2867 2980 KRINGLE TYPE IV, 26.
FT DOMAIN 2981 3094 KRINGLE TYPE IV, 27.
FT DOMAIN 3095 3208 KRINGLE TYPE IV, 28.
FT DOMAIN 3209 3322 KRINGLE TYPE IV, 29.
FT DOMAIN 3323 3436 KRINGLE TYPE IV, 30.
FT DOMAIN 3437 3550 KRINGLE TYPE IV, 31.
FT DOMAIN 3551 3664 KRINGLE TYPE IV, 32.
FT DOMAIN 3665 3770 KRINGLE TYPE IV, 33.
FT DOMAIN 3771 3884 KRINGLE TYPE IV, 34.
FT DOMAIN 3885 3998 KRINGLE TYPE IV, 35.
FT DOMAIN 3999 4112 KRINGLE TYPE IV, 36.
FT DOMAIN 4113 4226 KRINGLE TYPE IV, 37.
FT DOMAIN 4227 4327 KRINGLE TYPE V.
FT ACT_SITE 4328 4548 SERINE PROTEASE.
FT ACT_SITE 4369 4369 CHARGE RELAY SYSTEM.
FT ACT_SITE 4412 4412 CHARGE RELAY SYSTEM.
FT ACT_SITE 4498 4498 CHARGE RELAY SYSTEM.
FT VARIANT 4193 4193 W -> R (LOSS OF LYSINE-SEPHAROSE
FT BINDING).
FT FTID=VAR 006633.
SQ SEQUENCE 4548 AA; 501313 MW; 96921BE96A465C5F CRC64;

Query Match 21.2%; Score 478; DB 1; Length 4548;
Best Local Similarity 36.0%; Pred.No. 1.3e-29;
Matches 129; Conservative 36; Mismatches 133; Indels 60; Gaps 17;

Qy 50 CYEGNGHFYRGKASTDTMGKPLPWSATVLOQTYHAHRSALQLGLGK-----HNYCRN 104
Db 4228 CMFNGKGYRGKATTVGTGTCQEW-----AAQEPHRSHTFIPGNTKMWAGLEKKNYCRN 4280
Qy 105 PD-NRRRPWCYVQVGLKPLVQECWVHDCADGKLFQCGQKTLRPR---FKIIGBEFTIE 160
Db 4281 PDGIDNGPWCYT-WNPKLFDYCDIPLCASS--SFDGKQVPEFKPCPSIVGCGVAHPH 4337
Qy 161 NQFWFAIYRRHRGGSVTVVCGGLSPICWVISAHCFFIDYPKKEDYIVYLGSR---LN 217
Db 4338 SWPQVSL--RTRFGK--HFCGGTULISPENVLTAACHLKSSRPSSYKVLGAHQEWNL 4393

```

QY 218 SNTQGMKEFEVENLILHKDYSADTLAHNDIALLKIRSKGRCQPSRTIQICLPSMYN 277
DB 4394 SHVQ-----EIEVSLFLEPTQA-----DIALKL-----SRPAVITDKVPACLPFS--- 4435
QY 278 DPOF-----GTSCEITGFKENSTDYLYPQLKMTVVKLISHRECOQPHYGVSEVITKMLC 333
DB 4436 -PDYMTARTCEVITGWGTQT--FGTGLLKEAQLLVITENEVCN--HY-----KYIC 4483
QY 334 ADPQWKTSCQDSDGGLPVCSLQGRMTLTGIVSWGRGALKDKKPGVYTRVSHFLPWI 391
DB 4484 AEHLARGTSCQDSDGGLPVCFEKDKYILQGVTSWGLGCARPKNKPGVYARVSRFVTWI 4541

RESULT 29
APOA_MACMU STANDARD; PRT; 1420 AA.
AC P1417;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Apolipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Lp(a)) (fragment).
GN LPA.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89174660; PubMed=2925643;
RA Tomlinson J.E., McLean J.W., Lawn R.M.;
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
RT synthesis.";
RL J. Biol. Chem. 264:5957-5965(1989).
CC -I- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp 330.
CC -I- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC decorin (By similarity).
CC -I- PTM: N- and O-glycosylated (By similarity).
CC -I- DISEASE: Elevated plasma concentrations of apo(a) and its
CC naturally occurring proteolytic fragments are correlated with
CC atherosclerosis. Homology with plasminogen kringle IV and V is
CC thought to underlie the atherogenicity of the protein, because the
CC fragments are competing with plasminogen for fibrin(ogen) binding.
CC -I- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
CC leading to the formation of the so called mini-Lp(a). Apo(a)
CC fragments accumulate in atherosclerotic lesions, where they may
CC promote thrombogenesis. O-glycosylation may limit the extent of
CC proteolytic fragmentation (By similarity).
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -I- SIMILARITY: Contains at least 10 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J04635; AAA36833.1; -
DR PIR; A32869; A32869.
DR HSP; P00747; 2PK4.
DR MEROPS; S01.226; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle_1.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPsin.

```

```

DR PRINTS; P000018; KRINGLE.
DR ProDom; PD000395; Kringle; 10.
DR SMART; SMO0130; KR; 10.
DR PROSITE; SMO0020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 10.
DR PROSITE; PS00070; KRINGLE_2; 10.
DR PROSITE; PS0240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; FALSE_NEG.
DR PROSITE; PS00135; TRYPsin_SER; FALSE_NEG.
KW Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
KW Kringle; Repeat; Atherosclerosis.
FT NON_TER 1
FT DOMAIN 49 127 KRINGLE 1.
FT DOMAIN 163 241 KRINGLE 2.
FT DOMAIN 277 355 KRINGLE 3.
FT DOMAIN 391 469 KRINGLE 4.
FT DOMAIN 505 583 KRINGLE 5.
FT DOMAIN 619 697 KRINGLE 6.
FT DOMAIN 725 803 KRINGLE 7.
FT DOMAIN 839 917 KRINGLE 8.
FT DOMAIN 953 1031 KRINGLE 9.
FT DOMAIN 1067 1145 KRINGLE 10.
FT DOMAIN 1191 1420 SERINE PROTEASE.
SQ SEQUENCE 1420 AA; 158367 MW; BE102949E0JCS80E CRC64;

Query Match 21.1%; Score 477; DB 1; Length 1420;
Best Local Similarity 34.0%; Pred. No. 4.1e-30;
Matches 129; Conservative 43; Mismatches 137; Indels 70; Gaps 15;

QY 50 CYEGNGHYRGKASTDIMGRCPLPNSATVLQ--QTYAHRSDALQLGLGKHNYCRNPDN 107
DB 1068 CYHGNGQSYRGFTSTVTGRTCSQSWSMTPHQHKRTPENHPDMLTM-----NYCRNPD 1122
QY 108 RRRPWCYVQVGLKPLVQ--CMVHDCAD-----GKLKFCQG 142
DB 1123 DTGWCFT--WDPSVREYCNLFRCSDEGTCTVTPTVTPVPSLEAPSEQASSSFCGK 1179
QY 143 KTLPR---FKIIGEFFTIENQWFAIYRRHGGSVYVCGSLSPCWISATHCFI 199
DB 1180 PQVEFKPGSGIVGCVVAHPHSPWQSL--RTRFGK--HFCGGLTISPFWLTAACCLE 1235
QY 200 DYPKKEDIVVLGRSR---LNSQTQGMKEFEVENLILHKDYSADTLAHNDIALLKIRSK 256
DB 1236 TFSRPSFKVLGAHQEVNLESHVQ---EIEVSLFLEPTQA-----DIALKL--- 1281
QY 257 EGRCAPSRITQICLPSMYNDPQF---GTSCEITGFKENSTDYLYPEQLKMTVVKLI 312
DB 1282 -SRPAIITDKVIPACLPFS---PNVITAMTECVITGWGTQT--FGAGLLKEAQLHVI 1334
QY 313 SHRECOQPHYGVSEVITKMLCAADPQWKTSCQDSDGGLPVCSLQGRMTLTGIVSWGRGC 372
DB 1335 ENTVCNHYEFLNGRVKSTELCAGHLAGGTDRCCQDNGGPGVVCFDKDKYILRGITSWGPGC 1394
QY 373 ALKDKPGVYTRVSHFLPWI 391
DB 1395 ACPNKGYYVRSFVTWI 1413

RESULT 30
PLMN_HORSE
ID PLMN_HORSE STANDARD; PRT; 338 AA.
AC P80010;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Plasminogen (EC 3.4.21.7) (fragment).
GN PLG.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:23 ; Search time 49.5854 Seconds
(without alignments)
2097.294 Million cell updates/sec

Title: US-09-880-503-6
Perfect score: 2257
Sequence: 1 SNELHQVPNCCLNGGTCV.....VSHFLPWIRSHKTEENGLAL 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mnc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	1885	83.5	433	6	Q8MILO	oryctolagus
2	1880	83.3	433	6	Q8MHY7	oryctolagus
3	953.5	42.2	214	6	Q9XT70	oryctolagus
4	852.5	37.8	516	4	Q9BU99	homo sapien
5	824.5	36.5	559	11	Q91VP2	mus musculus
6	823	36.5	564	6	Q8MKB1	oryctolagus
7	815.5	36.1	562	6	Q8SQ23	mus musculus
8	808.5	35.8	231	11	Q8C612	mus musculus
9	788.5	34.9	395	4	Q9BZWI	homo sapien
10	778	34.5	154	4	Q9ES88	homo sapien
11	716	31.7	653	11	Q8VCS4	mus musculus
12	680.5	30.2	616	6	Q97507	mus musculus
13	678	30.0	615	4	Q81Z25	homo sapien
14	668	29.6	597	11	Q35727	mus musculus
15	664.5	29.4	517	11	Q8K0D2	mus musculus
16	651.5	28.9	560	4	Q14520	homo sapien

17	638	28.3	157	6	Q9TVAB	bos taurus
18	585	25.9	128	6	Q97587	oryctolagus
19	505	22.4	812	11	Q91WJ5	mus musculus
20	500	22.2	812	11	Q9ROW3	rattus norv
21	498	22.1	810	4	Q15146	homo sapien
22	495	21.9	103	6	Q95M89	equus caball
23	493.5	21.9	300	4	Q96EF3	homo sapien
24	493	21.8	454	6	Q46506	papio hamad
25	491.5	21.8	429	13	Q8AVB0	brachydanio
26	486.5	21.6	334	6	Q46507	papio hamad
27	469.5	20.8	806	6	Q18783	macropus eu
28	457	20.2	868	5	Q9Y1V3	polyandroca
29	445.5	19.7	327	4	Q8N171	homo sapien
30	438.5	19.4	284	4	Q8NF86	homo sapien
31	430.5	19.1	505	5	Q966V4	halocynthia
32	428	19.0	537	4	Q9BYE1	homo sapien
33	427.5	18.9	761	11	Q99JC8	rattus norv
34	426.5	18.9	267	5	Q9BK47	luidia foli
35	424	18.8	581	4	Q9BYE2	homo sapien
36	423	18.7	471	11	Q8CFE0	mus musculus
37	423	18.7	802	4	Q8TUE2	homo sapien
38	423	18.7	811	4	Q8IU80	homo sapien
39	415.5	18.4	799	11	Q9DB10	mus musculus
40	411.5	18.2	276	11	Q8CGR6	mus musculus
41	411.5	18.2	638	11	Q8R0P5	mus musculus
42	411	18.2	371	11	Q8CJ16	rattus norv
43	411	18.2	445	11	Q8CJ17	rattus norv
44	404	17.9	455	11	Q8CDR0	mus musculus
45	401.5	17.8	422	4	Q8WVC1	homo sapien

ALIGNMENTS

RESULT 1

ID	Q8MILO	PRELIMINARY;	PRT;	433 AA.
AC	Q8MILO;			
DT	01-OCT-2002 (TREMBlrel. 22, Created)			
DT	01-OCT-2002 (TREMBlrel. 22, Last sequence update)			
DT	01-MAR-2003 (TREMBlrel. 23, Last annotation update)			
DE	Urokinase-type plasminogen activator.			
GN	PLAU			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
-OX	NCBI_TaxID=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22155945; PubMed=12149463;			
RA	Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,			
RA	Dichek D.A.;			
RT	"Increased expression of urokinase during atherosclerotic lesion			
RT	development causes arterial constriction and lumen loss, and			
RT	accelerates lesion growth."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).			
CC	-1- SIMILARITY: BSLONGS TO PEPTIDASE FAMILY S1.			
CC	-1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.			
DR	EMBL; A122285; AAM83187.1; -			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR006209; EGF like.			
DR	InterPro; IPR000001; Kringle.			
DR	InterPro; IPR001254; Ser_protease_Try.			
DR	Pfam; PF00051; kringle; 1.			
DR	Pfam; PF00089; tryptsin; 1.			
DR	PRINTS; PRO0722; CHYMOTRYPSIN.			
DR	PRINTS; PRO0019; KRINGLE.			
DR	ProDom; PD000395; Kringle; 1.			
DR	SMART; SM00130; KR; 1.			
DR	SMART; SM00020; Tryp_spc; 1.			
DR	PROSITE; PS00022; EGF_1; 1.			
DR	PROSITE; PS00021; KRINGLE_1; 1.			
DR	PROSITE; PS00070; KRINGLE_2; 1.			

DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65E64F36415549B0 CRC64;

Query Match 83.5%; Score 1885; DB 6; Length 433;
Best Local Similarity 82.8%; Pred. No. 2e-172;
Matches 342; Conservative 23; Mismatches 39; Indels 10; Gaps 2;

OY 1 SNEHQV--PSNCDLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDSKTCYEGNGHY 58
DB 21 SHELHGVSDASNCGLNGGTCVTKYFSNIHWCNCPKFGQGHCEIDSKTCYEGNGHY 80
OY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVG 118
DB 81 RGKANTIDMRPCLAWNSANVLTQTYHAHHPDALQGLGKHNYCRNPDHRRPWCYVQVG 140
OY 119 LKPLVQECMVHDCAD-----GKLKFCQCKTLRPRFKIIGGEFTTIENQPFWFAAIYR 170
DB 141 LKQLIQECKVHDCSSGKKPALPPGKLEFQCGOKALRPRFKIIGGEFTTIENQPFWFAAIYR 200
OY 171 RHGGSVTVCGGSLISPCWVSAHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVEN 230
DB 201 RHGGSVTVCGGSLISPCWVSAHCFINHQKEDYIVLGRSLNSMTPGEMKFEVEQ 260
OY 231 LILHKDYSADTLAHNNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCITGF 290
DB 261 LILHEGYADTLAHNNDIALKILSNNGCAQPSRSITICLPPWADPNFGTSCITGF 320
OY 291 KGENSTLYPEQLKMTVVKLISHRECOQPHYGVSEVTTMLCAADPQWTKDSQGDSSG 350
DB 321 KGENSTLYPEQLKMTVVKLVSYQECQPHYGVSEVTTMLCAADPQWTKDSQGDSSG 380
OY 351 PLVCSQGRMTLTGIVSGRGCAKDKPGVTVRVSHFLPWIRSHIRSHIGEENGLAL 403
DB 381 PLVCSVQGRMTLTGIVSGRGCAKDKKPKGVTVRVSRFLPWIRSHIRSHIGEENGLAL 433

RESULT 2
Q8MHY7
ID Q8MHY7 PRELIMINARY; PRT; 433 AA.
AC Q8MHY7;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sug.ki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA, complete cds."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AV029517; AAK40239.1; -
DR EMBL; AB087224; BAC02685.1; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 6DD35A371010A6EE CRC64;

Query Match 83.3%; Score 1880; DB 6; Length 433;
Best Local Similarity 82.6%; Pred. No. 5.9e-172;
Matches 341; Conservative 23; Mismatches 39; Indels 10; Gaps 2;

OY 1 SNEHQV--PSNCDLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDSKTCYEGNGHY 58
DB 21 SHELHGVSDASNCGLNGGTCVTKYFSNIHWCNCPKFGQGHCEIDSKTCYEGNGHY 80
OY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVG 118
DB 81 RGKANTIDMRPCLAWNSANVLTQTYHAHHPDALQGLGKHNYCRNPDHRRPWCYVQVG 140
OY 119 LKPLVQECMVHDCAD-----GKLKFCQCKTLRPRFKIIGGEFTTIENQPFWFAAIYR 170
DB 141 LKQLIQECKVHDCSSGKKPALPPGKLEFQCGOKALRPRFKIIGGEFTTIENQPFWFAAIYR 200
OY 171 RHGGSVTVCGGSLISPCWVSAHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVEN 230
DB 201 RHGGSVTVCGGSLISPCWVSAHCFINHQKEDYIVLGRSLNSMTPGEMKFEVEQ 260
OY 231 LILHKDYSADTLAHNNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCITGF 290
DB 261 LILHEGYADTLAHNNDIALKILSNNGCAQPSRSITICLPPWADPNFGTSCITGF 320
OY 291 KGENSTLYPEQLKMTVVKLISHRECOQPHYGVSEVTTMLCAADPQWTKDSQGDSSG 350
DB 321 KGENSTLYPEQLKMTVVKLVSYQECQPHYGVSEVTTMLCAADPQWTKDSQGDSSG 380
OY 351 PLVCSQGRMTLTGIVSGRGCAKDKPGVTVRVSHFLPWIRSHIRSHIGEENGLAL 403
DB 381 PLVCSVQGRMTLTGIVSGRGCAKDKKPKGVTVRVSRFLPWIRSHIRSHIGEENGLAL 433

RESULT 3
Q9XT70
ID Q9XT70 PRELIMINARY; PRT; 214 AA.
AC Q9XT70;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator (fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Yin J., Idell S.;
RT "Partial mRNA of rabbit uPA."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097847; AAD39351.1; -
DR HSP; P00749; IEJN.
DR MEROPS; S01.231; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.

DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
 FT NON_TER 214
 SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;
 Query Match 42.2%; Score 953.5; DB 6; Length 214;
 Best Local Similarity 81.3%; Pred. No. 2.3e-83;
 Matches 174; Conservative 12; Mismatches 21; Indels 7; Gaps 1;
 QY 67 MGRPCLPWNATVLOQTVAHRSALQGLGKHNYCRNPNRRRCWCVQVGLKPLVQEC 126
 Db 1 MDRPCLANNSANVLTQYHAHPRDALQGLGKHNYCRNPNRRRCWCVQVGLKPLVQEC 60
 QY 127 MYHDC-----DGKLFQCCQKTLRPRFKIIGGEFTTIENOPFAAIYRRHGGSVTY 179
 Db 61 KVHDSGKXPALPPGKLEFCQCKLRPRFKIIGGEFTTIENOPFAAIYRRHGGSVTY 120
 QY 180 VCGGSLSPCWISATHCFIDYPKEDYIVLGRSLNSNTQEMKFEVENILHKDYSA 239
 Db 121 VCGGSLSPCWVVSATHCFINQKEDYIVLGRSLNSMTGEMKFEVEQLLHEGYRA 180
 QY 240 DTLAHNDIALKIRSKRGCAQPSRTIOTICLP 273
 Db 181 DTLAHNDIALKILSNNGCAQPSRSIOTICLP 214
 RESULT 4
 Q9BU99 PRELIMINARY; PRT; 516 AA.
 ID Q9BU99
 AC Q9BU99
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to plasminogen activator, tissue.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Skin.
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 CC EMBL; BC002795; AA02795.1; -
 DR HSP; P00750; IASH.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR006210; EGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00022; EGF 1; 1.
 DR PROSITE; PS01186; EGF 2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 2.
 DR PROSITE; PS00070; KRINGLE_2; 2.

DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KW Serine protease.
 SQ SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;
 Query Match 37.8%; Score 852.5; DB 4; Length 516;
 Best Local Similarity 37.1%; Pred. No. 3.5e-73;
 Matches 184; Conservative 55; Mismatches 154; Indels 103; Gaps 9;
 QY 1 SNELH-----QVPSNCDCLNGTCTVSNKYFNSNIHWCNPKKFGGHCHEIDKSKTC 50
 Db 23 SOEIHAFRRRGARSYQGCSEPRCFNGTCQALYFSDF--VCQPEGAGKCEIDTRATC 81
 QY 51 YEGNGHFVRGKASDTWGRCLPWNATVLOQTVAHRSALQGLGKHNYCRNPNRRR 110
 Db 82 YEDGIGSYRGTWSTAESGAECTWNSSALAQKPSGRPRDAIRLGLGNHNYCRNPNDRSK 141
 QY 111 PWCYVQVGLKPLVQECMVHPCADG----- 134
 Db 142 PWCYVFRAGKYSSEFCSTPACSEGNSDCYFGNSAYRGTHSLTESGASCLPWNMILIGK 201
 QY 135 -----KLKFO-----CG-OKTLRP 147
 Db 202 VYTAQNPAAQALGKGNKYNCRNPDGDAKPMCHVKNRRLTWECVDFSCCTGLRQYSQP 261
 QY 148 RFKIIGGEFTTIENOPFAAIYRRH-RGGSVYVYVCGSLISPCWVISAHCIFIDYPKED 206
 Db 262 QFRIGKGLFADIASHPWQAIFAXHRSRPGERFLCGGILISSCWLSAAHCFQERFPPHH 321
 QY 207 YIVYLGSRSLNSNTQEMKFEVENILHKDYSAADTLAHNDIALKIRSKRGCAQPSRT 266
 Db 322 LTVILGRTRYVVPGEEEKFEVEKYIVHKEFDDDT--YNDIALQLKLSDSRCAQESSV 379
 QY 267 IOTICLPSMYNDPQGTSCITGFKENSTDYLPQELKMTVWKLISHRECCQPHYYGSE 326
 Db 380 VRTVCLPPADQLPDWTECELSYGKHEALSPFYSERLKEAHVLYPSSRCTSQHLLNRT 439
 QY 327 VTRKMLCAAD-----PQWKT--DSCGDSGGPLVCSLQGRMTLTGIVSHVGRGCALKDKPGV 380
 Db 440 VTDNMLCAGDTRSGGPQANLHDACQDGGSGPLVCLNDGRMTLVGIISWGLCGGQKDVPGV 499
 QY 381 YTRUSHFLPWIRSHTK 396
 Db 500 YTKVTNYLOWIRDNR 515
 RESULT 5
 Q91VP2 PRELIMINARY; PRT; 559 AA.
 ID Q91VP2
 AC Q91VP2
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to plasminogen activator, tissue.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Breast tumor;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 DR EMBL; BC011256; AA011256.1; -
 DR HSP; P00761; IAN1.
 DR MGD; MGI:97610; Plat.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR006209; EGF like.


```
Db 431 VCLPDASLQLPDWTCELSGKHEFFSPVSEQLKEAHVRLYPSRCTPQOLKHTVTG 490
QY 330 KMLCAADPQW-----KTDSQQDSGGPLVCSLQGRMTLTGIVSWGRCALXDKPGVYTR 383
Db 491 NMLCAGDTRSGGAQVNLHDACQDSGGPLVCMTDGHTLTIGIISWGLGGQKDVPGVYTK 550
QY 384 VSHFLPWIRSH 394
Db 551 VVNYLGIQOH 561

RESULT 7
ID Q8SQ23 PRELIMINARY; PRT; 562 AA.
AC Q8SQ23;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9923;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Enamel organ;
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF364605; AA00297.1; -.
DR HSP; P00761; IAN1.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibnctn1.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS01186; EGF 2; 1.
DR PROSITE; PS01253; FIBRONECTIN 1; 1.
DR PROSITE; PS00021; KRINGLE 1; 2.
DR PROSITE; PS00070; KRINGLE 2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 562 AA; 63668 MW; F9E6B4C77CB101E8 CRC64;

Query Match 36.1%; Score 815.5; DB 6; Length 562;
Best Local Similarity 35.9%; Pred. No. 1.4e-69;
Matches 175; Conservative 60; Mismatches 156; Indels 97; Gaps 11;

QY 3 ELHQPV-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFY 58
Db 77 QCHSVPKVSCSEPCFNGGTCLOAIFSDP-VCCQPVGFGRQCEIDARATCYEDGDIY 135
QY 59 RGKASTDTMGRLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVG 118
Db 136 RGTWSTTESGAECVNMWNTSLASMPYNGRPPDAVKLGLGNHNYCRNPDKDKPWCYIFKA 195
QY 119 LK-----PLV-----QECWV----- 128
Db 196 EKYSDFCSTPACTKEKECYTGKGLDYGRTRSLTWSGAFCLPWNLSVLMGKIYTAWN 255
```

```
QY 129 -----HDCADGKLKFO-----CGQKTLR-PREKIIGE 155
Db 256 AQTGLGKHNYCRNPDGDTQWCHVLKDKHKLWTWCYDLPCQVTCGLRQYKEPQRIKGL 315
QY 156 FTTIENOPWFAAIY-RRHRGGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVYLGSR 214
Db 316 YADITSHPWQAIFVKNNRSPGERFLCGGILISSCWLSAAHCFQERPPPHHVRVLGRT 375
QY 215 RLNSNTQGEKMFVENLILHKDYSADTLAHNDIALLKIRSKGRCQAQPSITQITCLPS 274
Db 376 YRLVPGEEEAQFEVEKIVHKEFDDT--YDNDIALQLKSDSLTCAQESDAVRTVCLPE 433
QY 275 MYNDPQFGTSCITGFGKENSTDYLYPEOLKMTVVVKLIISHRECQOPHYVYGSEVTTKMLCA 334
Db 434 ANLQLPDWTECELSGKGHEASSPFYSERLKEAHVRLYPSRCTSKGLFNKTIINMLCA 493
QY 335 ADPQW-----KTDSQQDSGGPLVCSLQGRMTLTGIVSWGRCALXDKPGVYTRVSHFL 388
Db 494 GDRSGGDNANLHDACQDSGGPLVCMKNHMTLVGVISWGLGGQKDVPGVYTKVNTYL 553
QY 389 PWIRSHTK 396
Db 554 NWIRDNTR 561

RESULT 8
ID Q8C6L2 PRELIMINARY; PRT; 231 AA.
AC Q8C6L2;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Plasminogen activator (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
DR EMBL; AK054349; BAC35743.1; -.
FT NON_TER 231
SQ SEQUENCE 231 AA; 25510 MW; 25E8980A682737F2 CRC64;

Query Match 35.8%; Score 808.5; DB 11; Length 231;
Best Local Similarity 69.5%; Pred. No. 2.1e-69;
Matches 139; Conservative 21; Mismatches 31; Indels 9; Gaps 2;

QY 9 SNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRGKASTDTMG 68
Db 30 SNCGQNGGVCSYKYFSIRRCSCPRFQGEHCEIDASKTCYHNGSDSYGKANTDTKG 89
QY 69 RPLCPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLKPLVQECW 128
Db 90 RPLAWNAPVLQKPYNAHRPDAISLGLGKHNYCRNPDNKRKPWCYVQIGLRQFYQECW 149
QY 129 HDCADGKLK-----FCCGKTLRPFKLIIGEFTEIENQWPAATYRBRHGS-VTY 179
Db 150 HDCLSKSPSSSDVQQGFCQKALRPRFKLVGGFEVEENQWPAATYRBRHGSPPSF 209
QY 180 VCGGSLISPCWVISATHCFI 199
Db 210 KCGGSLISPCWVASAAHCFI 229

RESULT 9
```



```
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
CC -1- SIMILARITY: BELONGS TO EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAH19376.1; -.
DR HSSP; P00761; IANI.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00040; fn1; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00059; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match 31.7%; Score 716; DB 11; Length 653;
Best Local Similarity 36.3%; Pred. No. 66-60;
Matches 154; Conservative 58; Mismatches 162; Indels 50; Gaps 8;

QY 5 HQVPSNCDCLNGGTCVSNKYFSNIHW-----CNCPPKFGQGHCEIDKSKTCYEGNGH 56
DB 239 HTACLSSPCLNGGTC-----HLIVGTGTSVCTCPLGYAGRCNIVPTFHCFLNGT 289

QY 57 FYRGKASTDMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKYNCRNPDNRRRRCVYQ 116
DB 290 EYRGVASTAAGLSCLAWNSDLLYQLSHVDSVAAVLLGLGHAYCRNPKDQERPCVYV 349

QY 117 VGLKPLVQECMVHDC-----ADGKLKFCQGQK-----TLRPFKII 152
DB 350 KDNALSWECRLTACSLARVHSCSFEILALPESAPVRPTCGKHKRKTFLRPR--II 407

QY 153 GBFTTIIENQWPAALYYRRHGGSVTVVCGGSIISPCWISAPHCIFDIYPKKEDYIVYL 212
DB 408 GGSSSLPGSHPLAAIY---IGNS---FCAGSLVHTCWVVSAAHCFANSPRDSITVVLG 461

QY 213 RSLNSNTQCEMKFEVENLLHKDYSADTLAHNDIALKIRSEKRCQAPSETITICL 272
DB 462 QHFNRTTDTVQTGIEKVPYTYLVFNPNH-DLVLRKKKRGRCVRSQFVQFICL 520

QY 273 PSMYNDPQFGTSCEITGFGKENSTDVLYPEQLKMTVVVKLISHRECCQPHYGVSEVTKML 332
DB 521 PEAGSFTTGHKKQIAGWGHMDENVSSYNSLLLEALVPLVADHKCSSPEVYGADISPNWL 580
```

```
QY 333 CAADPQWKTDSCQDSDGGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTVRVSHFLPWIR 392
DB 581 CAGYFDCKSDACQDSDGGPLVCKRGVAYLYGIISWGGCGRLNKPQVYTVRVANVVDWIN 640

QY 393 SHTK 396
DB 641 DRIR 644

RESULT 12
O97507 PRELIMINARY; PRT; 616 AA.
AC O97507;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE FXII.
OS Sus scrofa (Fig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]_
RC TISSUE=Liver;
RA Takahashi T.; Kihara T.;
RT "Porcine liver factor XII.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AB022426; BAA37148.1; -.
DR HSSP; P00763; IDFO.
DR MEROPS; S01.211; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00059; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBBD1A9 CRC64;

Query Match 30.2%; Score 680.5; DB 6; Length 616;
Best Local Similarity 34.6%; Pred. No. 1.4e-56;
Matches 158; Conservative 59; Mismatches 156; Indels 83; Gaps 13;

QY 6 QVPSNCDCLNGGTCVSNKYFSNIHWNCPCPKFGQGHCEIDKSKTCYEGNGHFGKASTD 65
```

```
Db 176 QVCSNPLNGSGSLQTE---GHRLCRCPTGYAGRLCDVLDKERCYSVDRGLSVRGQAQIT 232
Qy 66 TMGRPCPLPWSNATVLOQTY-HAHRSDALQLGLKHNYCRNPNRRPFCWVQVGLKPLVQ 124
Db 233 LSGAPCQPMAS---EATYWNMTAEQALNGLGDHAFRCRPNDRTPFCVFWRGDQLSMQ 288
Qy 125 ECMVHDC---ADGKLKFO----- 139
Db 289 YCLRLARCOAIGEPAPILTQSPSEHQDPSLRSREPOPTTQPSQNLTSAMCAPPEQRG 348
Qy 140 -----CQKTLRPRF---KIIGFEFTTIENQWFAAIYRRHRGGSVTVVCGSLI 186
Db 349 PLPSAGLVGCCQR-LRRLSSLNRIVGLVALPAGFAPIAALYWGQ-----FCAGSLI 401
Qy 187 SPCWVISATHCFIDYPKKEDIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHN 246
Db 402 APCWLTAAACLQNRPAPELTVVLGQDRHNQSCQCTLAVRSYRLHESYSKTYQH-- 459
Qy 247 DIALLKIR-SKEGRCAQPSRTIQTICLP---SMYNDPQFGTSCEITGFGKENSTDVLYPE 302
Db 460 DLALVRLKETADGCCAHPSPFVQVCLPRSVASSAEPE-GALCEVAGHGQFEGAEYSS 518
Qy 303 QLKMTVVKLISHRECCQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPLVC---SIQGR 359
Db 519 FLQEAQVPLISPERCSAADVHGAAFTPGMLCAGFLEGGTDACQDGGSGPLVCEDETAERQ 578
Qy 360 MTLTGIVSWGRCALKDKPGVYTVRVSHFLPWRSH 395
Db 579 LVLRGIVSWGSGCDRLKPGVYTDVANYLAWIQEHT 614

RESULT 13
ID Q81Z25 PRELIMINARY; PRT; 615 AA.
AC Q81Z25;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Coagulation factor XII-Mie.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wada H., Nishioka J., Nakatani K., Kasai Y., Abe Y., Nobori T.;
RT "Molecular characterization of coagulation factor XII-Mie.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB095845; BAC23095.1; -.
SQ SEQUENCE 615 AA; 67735 MW; 030508870A0C7EDB CRC64;

Query Match 30.0%; Score 678; DB 4; Length 615;
Best Local Similarity 34.1%; Pred. No. 2.5e-56;
Matches 152; Conservative 57; Mismatches 159; Indels 78; Gaps 10;

Qy 13 CLNGTCVSNKYFSGNIHWCNCPKFGGCHCKIDSKTCYEGNGHFGYRGKASTDTMGRPCL 72
Db 183 CLHGGRCLE---VEGRLCHCPGVGTGPFCDVDTKASCYDGRGLSYRGLARTLLSGAPCQ 239
Qy 73 PWSNATVLOQTY-HAHRSDALQLGLKHNYCRNPNRRPFCWVQVGLKPLVQECMVHDC 131
Db 240 PMAS-----EATYRNVTAEQARNWGLGGHAFRCRPNDRIRPFCVLRNDRLSWEYCDLAQC 295
Qy 132 -----ADGKL----- 136
Db 296 QTPTQAAPTPVSRLVPLMPAQPAPKPPQPTTTPQSTTPGALPAKREQPPSLTRNG 355
Qy 137 KFOCCQ---KTLRPFKIIGFEFTTIENQWFAAIYRRHRGGSVTVVCGSLISPCWVIS 193
Db 356 PLSCQRLRKSLSSMTRVVGVALRGAPHAPIAALYWGHS-----FCAGSLIAPCWVLT 409
Qy 194 ATHCFIDYPKKEDIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKI 253
```

```
Db 410 AAHCLQDRPAPEDLTVVLGQERRHNSCBPCQTLAVRSYRLHAFS--PVSQHDLLALLRL 467
Qy 254 R-SKEGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDVLYPEQLKMTVVKLI 312
Db 468 QBDADGSCALLSPYQVPVCLPSGAARPSSETTLQCVAGGCHQFEGAEYASFLQEAQVPFL 527
Qy 313 SHRECCQPHYYGSEVTTKMLCAADPQWKTDSCQDGGSGPLVCSLQG---RMTLTGIVSWG 369
Db 528 SILERCSAPDVHGSSILPGMLCAGFLEGGTDACQDGGSGPLVCEQAAERRLLTQGIISWG 587
Qy 370 RCALKDKPGVYTVRVSHFLPWRSH 395
Db 588 SGCGRNKPVGYYTDVAYYLAWIREHT 613

RESULT 14
ID O35727 PRELIMINARY; PRT; 597 AA.
AC O35727;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Factor XII.
GN F12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue=Liver;
RC Schlessner M., Schwager S., Engel W.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; X9571; CAA67891.1; -.
DR HSSP; P00760; 1AQ7.
DR MEROPS; S01.211; -.
DR MGD; MGI:1891012; F12.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0013; FNTYPEII.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF_2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBROECTIN_1; 1.
DR PROSITE; PS00023; FIBROECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
```

```
SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0EBA CRC64;

Query Match 29.6%; Score 668; DB 11; Length 597;
Best Local Similarity 35.6%; Pred. No. 2.2e-55;
Matches 152; Conservative 61; Mismatches 154; Indels 60; Gaps 13;

QY 13 CLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRGKASTDTMGRPCL 72
DB 183 CLNGGSCD---LVEDHPLCRCPGTGYGFDLDLWATCYEGRLSVRGAGITQSGAPCQ 239
QY 73 PWSATVLOQTYHAHRSDALQLGLKHNCRPNRRRRCWVYVQVQ-----VQEC 126
DB 240 RW-----TVEATYRNMTKEQALSWGLGHAFRCRPNDRTPWCFCVMSGDRLSWDYCGLEQC 295
QY 119 -----LKLPLV-----QEVH-----DCADGKLKFG-----CGQ-----KTLRPRFKI 151
DB 296 QTPTFAPLVVPESGEESFSAPLSHAPNDSTHQTSLSTNTWGCQRFRKGLSGFMRV 355
QY 152 IGGSEFTTIENQWFAAIYRRHRGGSVTVVCGGLISPCWVISATHCFIDYPRKEDYIVYL 211
DB 356 VGLVALPGSHPIYALYWGNN-----FCAGSLIAPCWVLTAHCLQNRPAPEELTVVL 409
QY 212 GRSLNNTGEMKFEVENILHKDYSADTLAHNDIALILKIR-SKEGRCAQPSRTIOTI 270
DB 410 GQDRHNSCWCQTLAYRSVLHGFSSITYQH--DLALRLQESKTNSCALLSPHVQPV 467
QY 271 CLPMSYNDPQGTSCETITGFGKENSVDLYPEQLKMTWVKLISHRECCQPHYYGSEVTK 330
DB 468 CLPFGAAPPSETVLCEVAGHGHQLEGAEBYSTFLQEAQVFFIALDRCSNVHGDALLPG 527
QY 331 MLCAADPQWTKDSQCGSGGLPLVC---SLOGRMTLTGIVSWGRGCALKDPRGYTVRVSHF 387
DB 528 MLCAGFLEGGTDACQSGGLPLVCEGTAEHQTLRGVISMWGGCGDRNKPVGTVDVANY 587
QY 388 LPWIRSH 394
DB 588 LAWIQKH 594

RESULT 15
Q8K0D2 PRELIMINARY; PRT; 517 AA.
ID Q8K0D2;
AC Q8K0D2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC031775; AAH31775.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR001881; EGF Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000195; Kringle; 1.
DR SMART; SM00181; EGF; 3.

SMART; SM00130; KR; 1.
SMART; SM00020; Tryp_SPC; 1.
PROSITE; PS00022; EGF_1; 3.
PROSITE; PS01186; EGF_2; 2.
PROSITE; PS00021; KRINGLE_1; 1.
PROSITE; PS00070; KRINGLE_2; 1.
PROSITE; PS0240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 517 AA; 57326 MW; 3855A42035A5EA59 CRC64;

Query Match 29.4%; Score 664.5; DB 11; Length 517;
Best Local Similarity 36.9%; Pred. No. 3.9e-55;
Matches 154; Conservative 57; Mismatches 163; Indels 43; Gaps 14;

QY 13 CLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRGKASTDTMGRPCL 72
DB 116 CQGGVCGSRHRRSRF--TCACPDQYKGRFCEIGPD--DCVVGDSYGRGKSVKTVNQNPCL 173
QY 73 PWSATVLOQTYHAHRSDALQLGLKHNCRPNRRRRCWVYVQVGLKPL-----VQEC 126
DB 174 YNSHLLQETITNMPMEDAETHGIAEHNFCRNPDPGDHFKPWCFCVKNSEKVKWEYCDTVVC 233
QY 127 MYHDCADGKLKPT-----QCQKTLRPRF--KIIGSEFTTIENQWFAAIY---- 169
DB 234 PVPDTNPVSESLLEPVMELPGFESCGKTEVASHAVKRIYGGFKSTAGKHPQVQSLQSLP 293
QY 170 ---RRHRGGSVTVVCGGLISPCWVISATHCFIDYPRKEDYIVYLRSLNNTGEMKF 226
DB 294 LITSPQGG---HFCGGLIHPCWVLTAHC--TDINTKHLKVV--LGDQDLKKTESHEQTF 347
QY 227 EVENILHKDYSADTLAHNDIALILKIRSKGRCAQPSRTIOTICLPMSYNDP-QFGTSC 285
DB 348 RVEKILKYQYNERDEIPNDIALKLKXPVGCHCALESRYVKTVCPLS---DPFPGSTEC 404
QY 286 EITGFGKENSVDLYPEQLKMTWVKLISHRECCQPHYYGSEVTKMLCAADPQWK-TDSC 344
DB 405 HSGMGVTTETGE--GSRQLLDKAVKLIANPLCNSRLQYDHTTDDSMICAGNLQKPGSDTC 462
QY 345 QGDSGGPLVCSLQGRMTLTGIVSWGRGCALKDPRGYTVRVSHFPLPWIRSHTKENGL 401
DB 463 QGDSGGPLTCEKDGYYVYVYGVISWGQEGC--KKPGYVTVTKFLNWKITMTHREAGL 517

RESULT 16
Q14520 PRELIMINARY; PRT; 560 AA.
ID Q14520;
AC Q14520; Q00663;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE HGF activator like protein (Hyaluronan binding protein 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Kitamura N.;
RA Choi-Miura N.H.; Tobe T.; Sumiya J.; Nakano Y.; Sano Y.; Mazda T.;
RA Tomita M.;
RX MEDLINE=96425001; PubMed=8827452;
RA -Choi-Miura N.H.; Tobe T.; Sumiya J.; Nakano Y.; Sano Y.; Mazda T.;
RT "Purification and characterization of a novel hyaluronan-binding
RT protein (PHBP) from human plasma: it has three EGF, a kringle and a
RT serine protease domain, similar to hepatocyte growth factor
RT activator."
RL J. Biochem. 119:1157-1165(1996).
RN [3]
```

RP SEQUENCE FROM N.A.
RC TISSUE=Colon, and Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; D49742; BAA08576.1; -
DR EMBL; S83182; ABA46309.1; -
DR EMBL; BC031412; AAB11412.1; -
DR HSSP; P00763; IDPO.
DR MEROPS; S01.033; -
DR Genew; HGNC:4798; HAPB2.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 560 AA; 62671 MW; 5C1907230784ACD4 CRC64;

Query Match 28.9%; Score 651.5; DB 4; Length 560;
Best Local Similarity 36.4%; Pred. No. 7.7e-54;
Matches 152; Conservative 58; Mismatches 161; Indels 47; Gaps 14;

QY 13 CLNGGTCVSNKYFSNIHWCNCKPKFGGQHCIDKSKTCYEGNGHYRKGASTDTMGRPCL 72
DB 159 CONGATCSRHRKRSKF--TCACPDQPKGKFCBIG--SDDCYGDGGSYRGRKMTVYQNAACL 216
QY 73 PWSATVLTQOTYHAHRSALQGLGKHNYCRPNRRPWCYVQVGLKPLVQE--CMVHDC 131
DB 217 YWNSHLLIQENYMFMEADTHGIEHNFERNPDADKPKWCFIKVTNDKVKWECVDSAC 276
QY 132 ADGKUKP-----CQGQKTLPR--FKIIGGEFTTIEQPFALYRHR 173
DB 277 SAQDVAYPEESPTEPSTKLPGFDSGCKTEIAERKIKRIYGGFKSTAGKHPWQASLQ---- 332
QY 174 GGSVT-----YVCGGSLISPCWVISATHCFFIDYPKEDYIVVYGRSLNSNTQGEK 225
DB 333 -SSLPITSMQGHGFCGALHPCWVLTAAHC-TDI-KTRHLKVLDGQDLKKEEFHQ 389
QY 226 FEVENLIHKYSAATLAHNDIALKTRSGRCAQPSRTIQTICPSMYNDPQP--GT 283
DB 390 FRVEKIFKYSYNERDEPHNDIALLLKLPVDGHCALSKYKVTCLP-----DGSFSGS 445
QY 284 SCEITGFKENSTDVLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADQWK-TD 342
DB 446 ECHISGWGVTEGK--GSRQLDAVKLIANTLNSRQLYDHMDSDSMICAGN-QKPGQD 503
QY 343 SCQGDGGLVCSLQGRMTLTGIVSGRGCAKDKPKGVYTRVSHFLPWIRSHTEENG 400
DB 504 TCQGDGGLPTCEKQDGTGVYVYVGVSWGLECG--KRPGVYTVTKPLNWKATIKSESG 559

RESULT 17
Q9TVAS PRELIMINARY; PRT; 157 AA.
AC Q9TVAS8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=21071388; PubMed=11204721;
RA Balcerzak D., Querengesser L., Dixon W.T., Baracos V.E.;
RT "Coordinate expression of matrix-degrading proteinases and their
RT activators and inhibitors in bovine skeletal muscle.";
J. Anim. Sci. 79:94-107(2001).
RL J. Anim. Sci. 79:94-107(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF144761; AAD30301.1; -
DR HSSP; P00749; IURK.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR_1; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kinase; Kringle.
FT NON_TER 1
FT NON_TER 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7 CRC64;

Query Match 28.3%; Score 638; DB 6; Length 157;
Best Local Similarity 70.7%; Pred. No. 3.1e-53;
Matches 111; Conservative 14; Mismatches 24; Indels 8; Gaps 1;

QY 13 CLNGGTCVSNKYFSNIHWCNCKPKFGGQHCIDKSKTCYEGNGHYRKGASTDTMGRPCL 72
DB 1 CLNGGKCVTKYFSNTRQRCSPKFKGEGHEIDTSKTCYQNGHSGYRGNRDLGSRPCL 60
QY 73 PWSATVLTQOTYHAHRSALQGLGKHNYCRPNRRPWCYVQVGLKPLVQECMVHDC 132
DB 61 AWDSPTVLLKMYHAHRSALQGLGKHNYCRPNRRPWCYVQVGLKPLVQECMVHDC 120
QY 133 DGKL-----KFCGQKTLRPRFKIIGGEFTTIE 161
DB 121 VGKSPSPREKEEFQCGQKALRPRFKIVGGQVTAEN 157

RESULT 18
Q97587 PRELIMINARY; PRT; 128 AA.
AC Q97587;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand White;
RX MEDLINE=99057575; PubMed=9837780;
RA Reno C., Boykiw R., Martinez M.L., Hart D.A.;
RT "Temporal alterations in mRNA levels for proteinases and inhibitors
RT and their potential regulators in the healing medial collateral
RT ligament.";
J. Biochem. Biophys. Res. Commun. 252:757-763(1998).
RL Biochem. Biophys. Res. Commun. 252:757-763(1998).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF069711; AAC95003.1; -
DR HSSP; P00749; IELN.
DR MEROPS; S01.231; -

[illegible]


```

01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen activator urokinase (Fragment).
GN PLAU.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN _
SEQUENCE FROM N.A.
RP MEDLINE=21314992; PubMed=11421942;
RX Shubitsowski D.M., Venta P.J., Doughlass C.L., Zhou R.-X., Ewart S.L.;
RA "Polymorphism identification within 50 equine gene-specific sequence
RT tagged sites."
RL Anim. Genet. 32:78-78(2001)
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AY008806; AAK14840.1;
DR EMBL; AY008803; AAK14840.1; JOINED.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00089; trypsin; 1
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Hydrolase; Kinase; Protease; Serine protease.
FT NON_TER 1 1
FT TER 103 103
SQ SEQUENCE 103 AA; 11525 MW; 0B739514F6331180 CRC64;

Query Match 21.9%; Score 495; DB 6; Length 103;
Best Local Similarity 87.4%; Pred. No. 9.7e-40;
Matches 90; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 164 WFAIYRRHRGGSVTVVCGSLSPCWVLSATHCFIDYPKKEDYVYVGLGRSLNSNTQGE 223
DB 1 WFAIYRRHRGGSVTVVCGSLSPCWVLSATHCFINYPKEDYVYVGLGRSLNSSTSPGE 60
QY 224 MKFEVENLIHKDYSADTLAHNDIALLKIRSKRGCAQPSRT 266
DB 1 MKFEVENLIHKDYSADTLAHNDIALLKIRSKRGCAQPSRS 103

RESULT 23
Q96EF3 PRELIMINARY; PRT; 300 AA.
AC Q96EF3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN _
SEQUENCE FROM N.A.
RP TISSUE=Lung;
RA Strausberg R.;
RA Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; BC012390; AAH12390.1;
DR HSSP; P00761; IAN1
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; Hydrolase; Protease; Serine protease.
SQ SEQUENCE 300 AA; 32216 MW; FFC2BDF9382F536A CRC64;

01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen activator urokinase (Fragment).
GN PLAU.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN _
SEQUENCE FROM N.A.
RP MEDLINE=21314992; PubMed=11421942;
RX Shubitsowski D.M., Venta P.J., Doughlass C.L., Zhou R.-X., Ewart S.L.;
RA "Polymorphism identification within 50 equine gene-specific sequence
RT tagged sites."
RL Anim. Genet. 32:78-78(2001)
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AY008806; AAK14840.1;
DR EMBL; AY008803; AAK14840.1; JOINED.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00089; trypsin; 1
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Kinase; Protease; Serine protease.
FT NON_TER 1 1
FT TER 103 103
SQ SEQUENCE 103 AA; 11525 MW; 0B739514F6331180 CRC64;

Query Match 21.9%; Score 495; DB 6; Length 103;
Best Local Similarity 87.4%; Pred. No. 9.7e-40;
Matches 90; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 164 WFAIYRRHRGGSVTVVCGSLSPCWVLSATHCFIDYPKKEDYVYVGLGRSLNSNTQGE 223
DB 1 WFAIYRRHRGGSVTVVCGSLSPCWVLSATHCFINYPKEDYVYVGLGRSLNSSTSPGE 60
QY 224 MKFEVENLIHKDYSADTLAHNDIALLKIRSKRGCAQPSRT 266
DB 1 MKFEVENLIHKDYSADTLAHNDIALLKIRSKRGCAQPSRS 103

RESULT 24
Q46506 PRELIMINARY; PRT; 454 AA.
AC Q46506
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Apolipoprotein a (Fragment).
GN BABAP0A.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Papio.
OX NCBI_TaxID=9557;
RN _
SEQUENCE FROM N.A.
RA Cox L.A., Jett C., Hixson J.E.;
RA "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT Site Mutation is Associated with Deletion of a Single Exon in a Null
RT Allele."
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF029691; BAB97886.1;
DR HSSP; P00747; 2PK4.
DR MEROPS; S01.999;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Lipoprotein; Protease;
KW Serine protease.
FT NON_TER 1 1
SQ SEQUENCE 454 AA; 50041 MW; 974E30744C187B2F CRC64;

Query Match 21.8%; Score 493; DB 6; Length 454;
Best Local Similarity 33.8%; Pred. No. 9.8e-39;
```

Matches	127;	Conservative	51;	Mismatches	134;	Indels	64;	Gaps	14;								
QY	50	CYEGNGHGYRGKASDTMGRPCLPWN SATVLQ--QTYHAHRSDALQGLGKHNCRNPN	107														
Db	102	CYHGDCGSGYRGSEFTVTGRTCSGSSMTPHQHKRTPENHPNDGLTM-----NYCRNPD	156														
QY	108	RRRPWCYVQGLKPLV--QECMVHDCAD-----GKLKFCQCG	142														
Db	157	DTGCPWCF--MDPSVRWEYCNLTRGSDTEGTVVTLTVIPSLSEARSGQASSSFDCK	213														
QY	143	KTLRPR---FKIIGGFTTIENQWFAAIYRRHRGGSVTVVCGGSLSPCWVLSATHCTFI	199														
Db	214	PQVEPKKCPGRVVGCGVAHSHWPQVSL--RTRFGK--HFCGGTLLSPFWLTAARCLE	269														
QY	200	DYFKKEDIYVYGRSLNNTQCEMKFEVENLILHKDYSADTLAHNDIALLKIRSKGR	259														
Db	270	MSRPSPSYKVLG-----AHQEVNLSHVQIEVSKLFSPTGA---DIALKL-----SR	317														
QY	260	CAQPSRTIQICLPSMYNDPQF---GTCSEITGFGKENSTDYLPQLXMTVVKLISHR	315														
Db	318	PAITDKVIPACLPS---PNYVITATEYCIITGWGETGT--FGAGLLFEARLPVIENT	371														
QY	316	ECOPHYVGEVITKMLCAADPQWKTDSCOGDGGPLVCSLQGRMTLITGVSMGRGCALK	375														
Db	372	VCHRYEFLNGRVKSTELCAGHLAGGTDSCOGDGGPVVCFDKYILRGITSGPGCARP	431														
QY	376	DKPGVYTRVSHFLPWI	391														
Db	432	NKPGVYVRVSSFVTWI	447														
RESULT 25																	
Q8ABV0 PRELIMINARY; PRT; 429 AA.																	
AC	Q8ABV0																
DT	01-MAR-2003 (TrEMBLrel. 23, Created)																
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)																
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)																
DE	Plasminogen precursor (Fragment).																
OS	Brachydanio rerio (Zebrafish) (Danio rerio).																
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;																
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;																
OC	Cyprinidae; Danio.																
OC	NCBI_TaxID=7955;																
RN	[1]																
RN	SEQUENCE FROM N.A.																
RP	Hanumanthaiah R., Day K., Jagadeeswaran P.;																
RT	"Comprehensive analysis of blood coagulation pathways in teleostei.;																
RT	Evolution of coagulation factor genes and identification of zebrafish																
RT	factor VIII".																
RL	Blood Cells Mol. Dis. 0:0-0(2002).																
DR	EMBL; AF515276; AAN1006.1; --																
FT	NON TER																
SQ	SEQUENCE 429 AA; 47556 MW; 9A580A214A549C12 CRC64;																
Query Match 21.8%; Score 491.5; DB 13; Length 429;																	
Best Local Similarity 35.3%; Pred. No. 1.3e-38;																	
Matches 127; Conservative 37; Mismatches 155; Indels 41; Gaps 13;																	
QY	48	KTCYEGNGHGYRGKASDTMGRPCLPWN SATVLQ--TYHAHRSDALQGLGKHNCRN	104														
Db	100	KDCRNGNGAEYRGSTMTVMGVTCQWRSMTPHCHASFTPTHDPKGLG-----SNQCRN	154														
QY	105	PDNR--RRPWCYVQGLKPLVQECMVHDCADG															

Db 266 SKQERDVTKII---KGPAGT-----DIALLLK-----DRPALINDKVSFVCLPEKQYIVP 313

QY 281 FGTSCSITGFKENST---DYLYPEQLKMTVVKVLISHRECOQPHYVYGVSEVTTKMLCAADP 337

Db 314 SNTCEVVTGWGFEQDTGGEGY-----LKETGPFVIENKVCNRSFLNGRVKDHMCAGNI 368

QY 338 QWKTDSQCQSDSGPLVCSLQGRMTLGIYSWGRGCALKOKKPGVYTVRVSHFLPWIRSHTK 397

Db 369 EGGNDSQCQSDSGPLVCYAQNTFVLQGVTSWGLGCANAMKPGVYTVRVSKFVDWIERSIKE 428

RESULT 26

O46507 PRELIMINARY; PRT; 334 AA.

AC O46507; AC O46507; DB 6; Length 334;

DT 01-JUN-1998 (TrEMBLrel. 06, Created)

DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Plasmimogen (Fragment).

GN BABPSPSG.

OS Papio hamadryas (Hamadryas baboon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecoidea; Papio.

OX NCBI_TaxID=9557;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA Cox L.A., Jett C., Hixson J.E.;

RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice Site Mutation is Associated with Deletion of a Single Exon in a Null Allele."

RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR ENBL; AF029692; AAB97887.1; -.

DR HSP; P00747; SHPG.

DR MEROPS; S01.233; -.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003966; Prothrombin.

DR InterPro; IPR001254; Ser. protease_Try.

DR Pfam; PF00051; kringle; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR PRINTS; PR01505; PROTHROMBIN.

DR ProDom; PD000395; kringle; 1.

DR SMART; SM00130; KR; 1 SPC;

DR SMART; SM00020; TRY_P_SPC; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

DR PROSITE; PS0240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Glycoprotein, Hydrolase, Kringle; Protease; Serine protease.

FT NON-ITER 1

SQ SEQUENCE 334 AA; 36791 MW; C7DC06E03B965286 CRC64;

Query Match 21.6%; Score 486.5; DB 6; Length 334;

Best Local Similarity 36.2%; Pred. No. 2.8e-38;

Matches 129; Conservative 46; Mismatches 134; Indels 47; Gaps 16;

QY 50 CYEGNGHYRKGASTDTMGRCPLPWNSTVLTQTYAH-----RSDALQLGLGKHNYCRNP 105

Db 5 CMFNGNGRYRKGKATTVTGTCQBWAA-----KGFHSHLIETPTYPYRAGLEK-NYCRNP 58

QY 106 D-NRRRPMCVYQVGLKPLVQECMVHDGADGKLKFCQGOKTLRPR---PKIIGGEFTTIE 161

Db 59 DGDVGGPWCYT-TNPRKLYDCVDPQCASS--SPDCGKPEQVKKPCGRVVGCVAAH 115

QY 162 QPWFAIYRRHRGSGSVTVCGGSLISPCWVTSATHCFTDYPKKEDYIYVLGSRLSNNTQ 221

Db 116 WPQVSL--RTRFG--MHFCGGTILISPEWLVITAAHLEKSPRPSFYKVLGAHQ----- 165

```
QY 222 GEMKFE--VENLLHKDYSDTYLAHNDTALLKIRSEGRCAQPSRTIOTICLPSMYNDP 279
Db 166 -EVRLEPHVOEIEVSRKFSFAPGA---DIALLLSSP-----AIIKVIAPACLPSS---P 213
QY 280 QF-----GTSCEITGFGKENSNDLYPEQLKMTVVKLISHRECOQPHYYGVSEVTTKMLCAA 335
Db 214 NYVVADRTECFITGWGETQT--YGAGLLKEARLPVIENKVCNRYEFLNGRVKSTELCAG 271
QY 336 DPQWKTDSCGDSGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTVRVSHFPLWI 391
Db 272 HLAGGTDSCGDSGGPLVCFPEKDYILQGVTSWGLGCARPKNKPGVYTVRVSRFTWI 327

RESULT 27
Q18783 ID Q18783 PRELIMINARY; PRT; 806 AA.
AC Q18783;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Plasminogen
OS Macropus eugenii (Tamar wallaby)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9315;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=98004511; PubMed=9342350;
RA Law R.M., Schwartz K., Pethy L.;
RT "Convergent evolution of apolipoprotein(a) in primates and hedgehog.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997(1997).
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; AF012297; AAB65760.1; -.
DR HSP; P00747; SHPG.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 806 AA; 90981 MW; 95FAA86DC20064D5 CRC64;

Query Match 20.8%; Score 469.5; DB 6; Length 806;
Best Local Similarity 31.8%; Pred. No. 3.6e-36;
Matches 129; Conservative 43; Mismatches 165; Indels 67; Gaps 15;

QY 13 CLNGCTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFGYRGKASTDTMGPPCL 72
Db 448 CSGGSTVLNAQTRV-----PSVDTTSHPESD-----CMYSGSKDYRGKSTTTVTGLQC 498

QY 73 PWSATVLQOITYHAH---RSDALQLGLGKXNYCRNPD-NRRRPWCYVQVGLKPLVOECMV 128
Db 499 AWA-----QEPHRTTFTPTDYPRAGLEENYCNPNPDGNGPCWT-TNPKKLFQYCDI 552
```

```
QY 129 HDCADGKLKFOGQKTLRPR---FKIIGBEFTTIENQPFAAIYRRHRGGSVTVYCGGSL 185
Db 553 PCQVSPS-SFDCGRPRVEPKCPGRIVGCYAQPHSWPM--QISLRTFGE--HFCGGTL 607
QY 186 ISPCWVIATGCFIDYPKEDIYVYVGLSRSLNSNTQGMKFEVENLILHKDYSADTLAHH 245
Db 608 IAPQWVLTAACHLERSQWPGAYKVLG-----LHREYNPESYOE 647
QY 246 -----NDTALLKIRSEGRCAQPSRTIOTICLPSMYNDPQFGTSCIEITGFGKE 293
Db 648 TVSRLFGPLAADIALLLK-----NRPAINDKVIPACLPQDFMVPDRTLCHVTGWGDT 703
QY 294 NSTDYLYPE-QLKMTVVKLISHRECOQPHYYGVSEVTTKMLCAADPQWKTDCQGSGL 352
Db 704 QGTS---PRGLLKQASLPVIDNRVNCNRYEFLNGRVKSTELCAGHLVGRGDSQGSGL 760
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRVSHFPLWIRSHTK 396
Db 761 ICFEDDKYVLOGVTSWGLGCARPKNKPGVYTVRVSRFTWI 804

RESULT 28
Q9Y1V3 ID Q9Y1V3 PRELIMINARY; PRT; 868 AA.
AC Q9Y1V3;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Tunicate retinoic acid-inducible modular protease precursor.
GN TRAMP.
OS Polyandrocarpa misakiensis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Styelidae; Polyandrocarpa.
OX NCBI_TaxID=7723;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=white spot;
RX MEDLINE=99423846; PubMed=10491255;
RA Ohashi M., Kawamura K., Fujii N., Yubisui T., Fujiwara S.;
RT "A retinoic acid-inducible modular protease in budding ascidians.";
RL Dev. Biol. 214:38-45(1999).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AB030007; BAA82522.1; -.
DR HSP; P00763; LDPO.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Ser. protease_Try.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00057; ldl_recept_a; 3.
DR Pfam; PF00530; SRCR; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00261; LDLRECEPTOR.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00192; LDLa; 3.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00202; SR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS01209; LDLRA_1; 3.
DR PROSITE; PS00068; LDLRA_2; 3.
DR PROSITE; PS00287; SRCR_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease; Signal.
FT SIGNAL 1 28
SQ SEQUENCE 868 AA; 97660 MW; F71462865F36A6CA CRC64;

Query Match 20.2%; Score 457; DB 5; Length 868;
Best Local Similarity 39.2%; Pred. No. 6.3e-35;
Matches 104; Conservative 43; Mismatches 100; Indels 18; Gaps 9;
```

```
Qy 139 QCGKTL-----RRPKIIGSEFTTIENQPFAPAIYRRHRGGSVTVVCGSLISPCWVISA 194
Db 608 ECGKPVIEAPLPTARIVGSGTEPHEWPAQAGIWL-----PWTWCGGSLIHPCWVLT 662
Qy 195 THCFI-DYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKI 253
Db 663 AHCFVREYPR-DYIRLGDHITGVDDTEQLFKIAEIIKH-DYNVT--KENDIALLRI 718
Qy 254 RSKEGRCAQPSRTIQTICLPSMYNDPQGTSCETGFGKENSTDY-LYPQLKMTVVKLI 312
Db 719 ENDARECATITPEVQTVCLPKSSQFDAKTICEVTGWKDSATAYRAYVPLVQEAETPLI 778
Qy 313 SHRSCQPHYVGSVTTKMLCAADPQWKTDSQCGSLVCSLQ--RMTLGTIVSWGR 370
Db 779 ANKCLDSEY-TQLGTFMPCAGVLTGKDSQCGSLVCSLQ--RMTLGTIVSWGR 837
Qy 371 GCALKDKPGVYTRYVSHFELPWIRSH 395
Db 838 GCAKPKAPGVYAKVAFIDWIEQMT 862

RESULT 29
Q8N171 PRELIMINARY; PRT; 327 AA.
AC Q8N171;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Similar to protease, serine, 8 (Prostasin) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Brain;
RA Strausberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; BC036846; AAH36846.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1
SQ SEQUENCE 327 AA; 35106 MW; 60458ED817AC1CF3 CRC64;

Query Match 19.7%; Score 445.5; DB 4; Length 327;
Best Local Similarity 38.3%; Pred. No. 2.4e-34;
Matches 105; Conservative 36; Mismatches 100; Indels 33; Gaps 8;
Qy 134 GKLFQCGQKTLPRFKIIGSEFTTIENQPFAPAIYRRHRGGSVTVVCGSLISPCWVISA 193
Db 69 GKSAACQGPMSR--IVGGRDGRDGEWPQASI--QHRG---AHVCGSLIAPQWVLT 121
Qy 194 ATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKI 253
Db 122 AAHCFPRRALPAEYRVRLGALRLGSTSPRTLSVPVRRVLLPPDYSED--GARGDLALLQL 179
Qy 254 RSKEGRCAQPSRTIQTICLPSMYNDPQGTSCETGFGKENSTDYLP-----EQKLM 306
Db 180 R----RPVPLSARVQVCLPVFGARPPPGTCSRVTGWS-----LRPGVPLPEWRPQQ 229
Qy 307 TVVKLISHRECQPHYVGSVETTK-----MLCAADPQWKTDSQCGSLVCSLQGR 359
Db 230 VRVPLDLSRTCDGLYHVGADVPQAEIRIVLPGLSCAGYPQGHKDACQDGGPLTCLQSGS 289
Qy 360 MTLTGIVSWGRGKALKDKPGVYTRYVSHFELPWIRS 393

Query Match 19.4%; Score 438.5; DB 4; Length 284;
Best Local Similarity 38.0%; Pred. No. 9.3e-34;
Matches 104; Conservative 36; Mismatches 101; Indels 33; Gaps 8;
Qy 134 GKLFQCGQKTLPRFKIIGSEFTTIENQPFAPAIYRRHRGGSVTVVCGSLISPCWVISA 193
Db 22 GKSAACQGPMSR--IVGGRDGRDGEWPQASI--QHPG---AHVCGSLIAPQWVLT 74
Qy 194 ATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKI 253
Db 75 AAHCFPRRALPAEYRVRLGALRLGSTSPRTLSVPVRRVLLPPDYSED--GARGDLALLQL 132
Qy 254 RSKEGRCAQPSRTIQTICLPSMYNDPQGTSCETGFGKENSTDYLP-----EQKLM 306
Db 133 R----RPVPLSARVQVCLPVFGARPPPGTCSRVTGWS-----LRPGVPLPEWRPQQ 182
Qy 307 TVVKLISHRECQPHYVGSVETTK-----MLCAADPQWKTDSQCGSLVCSLQGR 359
Db 183 VRVPLDLSRTCDGLYHVGADVPQAEIRIVLPGLSCAGYPQGHKDACQDGGPLTCLQSGS 242
Qy 360 MTLTGIVSWGRGKALKDKPGVYTRYVSHFELPWIRS 393
Db 243 WVLGVVSWGRGKALKPNRPGVYTRYVSHFELPWIRS 276

Search completed: December 3, 2003, 14:42:59
Job time : 51.5854 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:43:04 ; Search time 41.3738 Seconds
(without alignments)
1811.566 Million cell updates/sec

Title: US-09-880-503-6
Perfect score: 2257
Sequence: 1 SNELHQPVSNCCLNGGTCV.....VSHFLPWIRSHKTBENGIAL 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2257	100.0	403	10	US-09-880-503-6
2	2243	99.4	411	10	US-09-880-503-3
3	2243	99.4	431	12	US-10-301-822-161
4	2243	99.4	431	12	US-10-131-985-21
5	2243	99.4	431	14	US-10-076-421-2
6	2243	99.4	431	15	US-10-171-311-184
7	2241	99.3	431	12	US-10-247-671-149
8	2240	99.2	431	15	US-10-193-656-4
9	2233	98.9	411	12	US-10-407-821-2
10	2225	98.6	431	10	US-09-264-4688-1
11	1703	75.5	323	10	US-09-880-503-7
12	1477	65.4	337	15	US-10-106-698-6266
13	1465	64.9	276	10	US-09-880-503-5
14	1464	64.9	268	12	US-10-407-821-3
15	1333	59.1	246	10	US-09-264-4688-2

16	1318	58.4	241	11	US-09-898-837A-47	Sequence 47, Appl
17	858.5	38.0	527	11	US-09-987-457-18	Sequence 18, Appl
18	858.5	38.0	527	11	US-09-987-455-19	Sequence 19, Appl
19	858.5	38.0	562	9	US-09-969-271-7	Sequence 7, Appl
20	858.5	38.0	562	10	US-09-974-298-145	Sequence 145, App
21	858.5	38.0	562	12	US-10-443-701-4	Sequence 4, Appl
22	858.5	38.0	562	15	US-10-193-656-8	Sequence 8, Appl
23	793	35.1	135	10	US-09-880-503-4	Sequence 4, Appl
24	793	35.1	138	10	US-09-984-186-12	Sequence 12, Appl
25	793	35.1	138	15	US-10-237-667-12	Sequence 12, Appl
26	793	35.1	138	15	US-10-237-708-12	Sequence 12, Appl
27	793	35.1	138	15	US-10-237-866-12	Sequence 12, Appl
28	793	35.1	138	15	US-10-237-871-12	Sequence 12, Appl
29	793	35.1	138	15	US-10-237-624-12	Sequence 12, Appl
30	793	35.1	143	10	US-09-880-503-8	Sequence 8, Appl
31	789.5	35.0	372	9	US-09-084-491A-3	Sequence 3, Appl
32	789.5	35.0	372	14	US-10-102-704-3	Sequence 3, Appl
33	784	34.7	354	11	US-09-987-457-10	Sequence 10, Appl
34	784	34.7	354	11	US-09-987-455-11	Sequence 11, Appl
35	784	34.7	377	11	US-09-987-457-14	Sequence 14, Appl
36	745	33.0	343	11	US-09-987-457-15	Sequence 15, Appl
37	745	33.0	343	11	US-09-987-457-12	Sequence 12, Appl
38	744	33.0	339	11	US-09-987-455-13	Sequence 13, Appl
39	744	33.0	339	11	US-09-987-457-11	Sequence 11, Appl
40	741	32.8	331	11	US-09-987-455-12	Sequence 12, Appl
41	741	32.8	331	11	US-09-987-457-15	Sequence 15, Appl
42	735	32.6	343	11	US-09-987-455-16	Sequence 16, Appl
43	735	32.6	343	11	US-09-987-457-13	Sequence 13, Appl
44	731	32.4	335	11	US-09-987-457-14	Sequence 14, Appl
45	731	32.4	335	11	US-09-987-455-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-880-503-6
; Sequence 6, Application US/09880503
; Patent No. US2002031964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-6

Query Match	100.0%	Score	2257	DB	10	Length	403
Best Local Similarity	100.0%	Pred. No.	1.4e-190				
Matches	403	Conservative	0	Mismatches	0	Indels	0
						Gaps	0
Qy	1	SNELHQPVSNCCLNGGTCVSNKYFSNIHWNCNPKFGQHCIEDKSKTCYEGNGHFYRG	60				
Db	1	SNELHQPVSNCCLNGGTCVSNKYFSNIHWNCNPKFGQHCIEDKSKTCYEGNGHFYRG	60				
Qy	61	KASDTMTGRCLPWSATVLOQTYHAHRSDALQGLGKHNYCRPNDRRPMCVYQVGLK	120				
Db	61	KASDTMTGRCLPWSATVLOQTYHAHRSDALQGLGKHNYCRPNDRRPMCVYQVGLK	120				
Qy	121	PLVQECMWHDCADGKLFQCGQKTLRPFKLIIGFEFTTIENQPMFAAIYRHRGGSVTV	180				
Db	121	PLVQECMWHDCADGKLFQCGQKTLRPFKLIIGFEFTTIENQPMFAAIYRHRGGSVTV	180				

Qy 181 CGGSLSPCWVISATHCFIDYPKKEDYVYVGLSRSLNSNTQGMKFEVENLIHKDYSAD 240
Db 181 CGGSLSPCWVISATHCFIDYPKKEDYVYVGLSRSLNSNTQGMKFEVENLIHKDYSAD 240
Qy 241 TLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPOFGTSCITGFGKNSDYLY 300
Db 241 TLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPOFGTSCITGFGKNSDYLY 300
Qy 301 PEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPOWKTDSCQDSGGPLVCSLQGRM 360
Db 301 PEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPOWKTDSCQDSGGPLVCSLQGRM 360
Qy 361 TLTGIVSGRCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 403
Db 361 TLTGIVSGRCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 403

RESULT 2

US-09-880-503-3
; Sequence 3, Application US/09880503
; Patent No. US2002031964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Root
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-3

Query Match 99.4%; Score 2243; DB 10; Length 411;
Best Local Similarity 98.1%; Pred. No. 2.5e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHYRG 60
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHYRG 60
Qy 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRRCWCVQVGLK 120
Db 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRRCWCVQVGLK 120
Qy 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
Db 121 PLVQECMVHDCADGKSPPEELAFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 180
Qy 173 RGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYVYVGLSRSLNSNTQGMKFEVENLI 232
Db 181 RGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYVYVGLSRSLNSNTQGMKFEVENLI 240
Qy 233 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPOFGTSCITGFGK 292
Db 241 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPOFGTSCITGFGK 300
Qy 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPOWKTDSCQDSGGPL 352
Db 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPOWKTDSCQDSGGPL 360
Qy 353 VCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 403
Db 361 VCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 411

RESULT 3

US-10-301-822-161
; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamakur, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029PZRN
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/391,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-161

Query Match 99.4%; Score 2243; DB 12; Length 431;
Best Local Similarity 98.1%; Pred. No. 2.6e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHYRG 80
Qy 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRRCWCVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRRCWCVQVGLK 140
Qy 121 PLVQECMVHDCADGK-----LKFCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172
Db 141 PLVQECMVHDCADGKSPPEELKFCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 200
Qy 173 RGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYVYVGLSRSLNSNTQGMKFEVENLI 232
Db 201 RGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYVYVGLSRSLNSNTQGMKFEVENLI 260
Qy 233 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPOFGTSCITGFGK 292
Db 261 LHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQTICLPSMYNDPOFGTSCITGFGK 320
Qy 293 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPOWKTDSCQDSGGPL 352
Db 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSVTTKMLCAADPOWKTDSCQDSGGPL 380
Qy 353 VCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 403
Db 381 VCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 431

RESULT 4

US-10-131-985-21
; Sequence 21, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V

APPLICANT: Huggins, Jonathan P
APPLICANT: McIntosh, Fraser S
APPLICANT: Occleston, Nicholas L
TITLE OF INVENTION: Composition
FILE REFERENCE: PCS 10391A
CURRENT FILING DATE: 2002-04-25
CURRENT FILING DATE: 2002-04-25
PRIOR FILING DATE: 2000-11-30
PRIOR FILING DATE: 2000-11-30
PRIOR FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-985-21

Query Match 99.4%; Score 2243; DB 12; Length 431;
Best Local Similarity 98.1%; Pred. No. 2.6e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQPFPAIYRRH 172
DB 141 PLVQECMVHDCADGKPPPEELKFCQCGKTLRPRFKIIGGEFTTIENQPFPAIYRRH 200
QY 173 RGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 232
DB 201 RGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCIEITGFGK 292
DB 261 LHKDYSADTLAHHNDIALKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCIEITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGVSEVTTMMLCAADPQWKTDSCQDSSGGPL 352
DB 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGVSEVTTMMLCAADPQWKTDSCQDSSGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 5
US-10-076-421-2
Sequence 2, Application US/10076421
Publication No. US20020193304A1
GENERAL INFORMATION:
APPLICANT: WADA, MANABU
APPLICANT: WADA, NAKO
TITLE OF INVENTION: ANTI-HIV AGENTS
FILE REFERENCE: HAYAK-9
CURRENT FILING DATE: 2002-05-17
PRIOR FILING DATE: 2001-02-20
PRIOR FILING DATE: 2001-06-19
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens

Query Match 99.4%; Score 2243; DB 12; Length 431;
Best Local Similarity 98.1%; Pred. No. 2.6e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQPFPAIYRRH 172
DB 141 PLVQECMVHDCADGKPPPEELKFCQCGKTLRPRFKIIGGEFTTIENQPFPAIYRRH 200
QY 173 RGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 232
DB 201 RGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCIEITGFGK 292
DB 261 LHKDYSADTLAHHNDIALKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCIEITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGVSEVTTMMLCAADPQWKTDSCQDSSGGPL 352
DB 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGVSEVTTMMLCAADPQWKTDSCQDSSGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

US-10-076-421-2

Query Match 99.4%; Score 2243; DB 14; Length 431;
Best Local Similarity 98.1%; Pred. No. 2.6e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK-----LKFCQCGKTLRPRFKIIGGEFTTIENQPFPAIYRRH 172
DB 141 PLVQECMVHDCADGKPPPEELKFCQCGKTLRPRFKIIGGEFTTIENQPFPAIYRRH 200
QY 173 RGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 232
DB 201 RGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 260
QY 233 LHKDYSADTLAHHNDIALKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCIEITGFGK 292
DB 261 LHKDYSADTLAHHNDIALKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCIEITGFGK 320
QY 293 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGVSEVTTMMLCAADPQWKTDSCQDSSGGPL 352
DB 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGVSEVTTMMLCAADPQWKTDSCQDSSGGPL 380
QY 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 6
US-10-171-311-184
Sequence 184, Application US/10171311
Publication No. US20030087270A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Chen, Yan
APPLICANT: Zhao, Xumei
APPLICANT: Monahan, John
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Giatt, Karen
APPLICANT: Gannavarapu, Manjula
APPLICANT: Hoerssh, Sebastian
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
TITLE OF INVENTION: OF CERVICAL CANCER
FILE REFERENCE: MRI-035
CURRENT APPLICATION NUMBER: US/10/171,311
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: US 60/298,159
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,155
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/335,936
NUMBER OF SEQ ID NOS: 238
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 184
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
US-10-171-311-184

Query Match 99.4%; Score 2243; DB 15; Length 431;
Best Local Similarity 98.1%; Pred. No. 2.6e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60

```
Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 80
Qy 61 KASTDTMGRPCLPWN SATVLQOYTHAHRSDALQGLGKHNYCRNPNRRPWCYVOVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOYTHAHRSDALQGLGKHNYCRNPNRRPWCYVOVGLK 140
Qy 121 PLVQECMVHDCADGK-----LKFCQGOKTLRPRFKIIGGEFTTIENQPFPAALYRRH 172
Db 141 PLVQECMVHDCADGKPPSPPEELKFQCGOKTLRPRFKIIGGEFTTIENQPFPAALYRRH 200
Qy 173 RGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFVENLI 232
Db 201 RGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFVENLI 260
Qy 233 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
Db 261 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
Qy 293 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 352
Db 321 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 380
Qy 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 7
US-10-247-671-149
; Sequence 149, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1
US-10-247-671-149

Query Match 99.3%; Score 2241; DB 12; Length 431;
Best Local Similarity 97.8%; Pred. No. 3.9e-189;
Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

Qy 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 80
Qy 121 PLVQECMVHDCADGK-----LKFCQGOKTLRPRFKIIGGEFTTIENQPFPAALYRRH 172
Db 141 PLVQECMVHDCADGKPPSPPEELKFQCGOKTLRPRFKIIGGEFTTIENQPFPAALYRRH 200
Qy 173 RGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFVENLI 232
Db 201 RGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFVENLI 260
Qy 233 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
Db 261 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
Qy 293 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 352
Db 321 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 380
Qy 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431
```

```
Qy 233 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
Db 261 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
Qy 293 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 352
Db 321 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 380
Qy 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 8
US-10-193-656-4
; Sequence 4, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: LI, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/IJ577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00749
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: (1)...(431)
US-10-193-656-4

Query Match 99.2%; Score 2240; DB 15; Length 431;
Best Local Similarity 97.8%; Pred. No. 4.8e-189;
Matches 402; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

Qy 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 80
Qy 61 KASTDTMGRPCLPWN SATVLQOYTHAHRSDALQGLGKHNYCRNPNRRPWCYVOVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOYTHAHRSDALQGLGKHNYCRNPNRRPWCYVOVGLK 140
Qy 121 PLVQECMVHDCADGK-----LKFCQGOKTLRPRFKIIGGEFTTIENQPFPAALYRRH 172
Db 141 PLVQECMVHDCADGKPPSPPEELKFQCGOKTLRPRFKIIGGEFTTIENQPFPAALYRRH 200
Qy 173 RGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFVENLI 232
Db 201 RGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFVENLI 260
Qy 233 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 292
Db 261 LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
Qy 293 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 352
Db 321 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSSGGPL 380
Qy 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403
```

Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 9

US-10-407-821-2

Sequence 2, Application US/10407821

Publication No. US20030219386A1

GENERAL INFORMATION:

APPLICANT: IDELL, STEVEN

TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED

TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL

TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS

FILE REFERENCE: UTSN-022US

CURRENT APPLICATION NUMBER: US/10/407,821

CURRENT FILING DATE: 2003-04-04

PRIOR APPLICATION NUMBER: 60/414,202

PRIOR FILING DATE: 2002-09-27

PRIOR APPLICATION NUMBER: 60/370,466

PRIOR FILING DATE: 2002-04-05

NUMBER OF SEQ ID NOS: 3

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 411

TYPE: PRT

ORGANISM: Homo sapiens

US-10-407-821-2

Query Match 98.9%; Score 2233; DB 12; Length 411;

Best Local Similarity 97.8%; Pred. No. 1.9e-188;

Matches 402; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

Qy 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFIYRG 60

Db 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFIYRG 60

Qy 61 KASDTMTGRPCLPNWSATVLOOTYHAHRSALQGLGKHNYCRPNRRPWCYVQVGLK 120

Db 61 KASDTMTGRPCLPNWSATVLOOTYHAHRSALQGLGKHNYCRPNRRPWCYVQVGLK 120

Qy 121 PLVQECMVHDCADGK-----LKFCQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 172

Db 121 LLVQECMVHDCADGKPPSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 180

Qy 173 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQEMKFEVENLI 232

Db 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQEMKFEVENLI 240

Qy 233 LHKDYSADTLAHHNDIALLKIRSKGRCAPSRITQITICLPSMYNDPQFGTSCIEITGFGK 292

Db 241 LHKDYSADTLAHHNDIALLKIRSKGRCAPSRITQITICLPSMYNDPQFGTSCIEITGFGK 300

Qy 293 ENSTDYLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDSSGGPL 352

Db 301 ENSTDYLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDSSGGPL 360

Qy 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403

Db 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 10

US-09-264-468B-1

Sequence 1, Application US/09264468B

Patent No. US20020106775A1

GENERAL INFORMATION:

APPLICANT: Wang, Jieyi

APPLICANT: Nienaber, Vicki L.

APPLICANT: Henkin, Jack

APPLICANT: Smith, Richard A.

APPLICANT: Walter, Karl A.

APPLICANT: Severin, Jean M.

APPLICANT: Egalji, Rohinton

APPLICANT: Johnson Jr., Robert W.

APPLICANT: Holzman, Thomas F.

TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE

FILE REFERENCE: 6310.US.P1

CURRENT APPLICATION NUMBER: US/09/264,468B

PRIOR FILING DATE: 1999-03-05

PRIOR APPLICATION NUMBER: US 09/036,361

PRIOR FILING DATE: 1998-03-06

NUMBER OF SEQ ID NOS: 23

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 431

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SIGNAL

LOCATION: (1)...(20)

OTHER INFORMATION: Leader sequence

NAME/KEY: VARIANT

LOCATION: (279)...(279)

OTHER INFORMATION: Xaa = any amino acid

NAME/KEY: VARIANT

LOCATION: (302)...(302)

OTHER INFORMATION: Xaa = any amino acid

US-09-264-468B-1

Query Match 98.6%; Score 2225; DB 10; Length 431;

Best Local Similarity 97.6%; Pred. No. 1e-187;

Matches 401; Conservative 0; Mismatches 2; Indels 8; Gaps 1;

Qy 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFIYRG 60

Db 21 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFIYRG 80

Qy 61 KASDTMTGRPCLPNWSATVLOOTYHAHRSALQGLGKHNYCRPNRRPWCYVQVGLK 120

Db 81 KASDTMTGRPCLPNWSATVLOOTYHAHRSALQGLGKHNYCRPNRRPWCYVQVGLK 140

Qy 121 PLVQECMVHDCADGK-----LKFCQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 172

Db 141 PLVQECMVHDCADGKPPSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 200

Qy 173 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQEMKFEVENLI 232

Db 201 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQEMKFEVENLI 260

Qy 233 LHKDYSADTLAHHNDIALLKIRSKGRCAPSRITQITICLPSMYNDPQFGTSCIEITGFGK 292

Db 261 LHKDYSADTLAHHNDIALLKIRSKGRCAPSRITQITICLPSMYNDPQFGTSCIEITGFGK 320

Qy 293 ENSTDYLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDSSGGPL 352

Db 321 EXSTDYLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDSSGGPL 380

Qy 353 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403

Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 11

US-09-880-503-7

Sequence 7, Application US/09880503

Patent No. US20020131964A1

GENERAL INFORMATION:

APPLICANT: CINES, Douglas B

APPLICANT: HIGAZI, Abd Al-Roof

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND

TITLE OF INVENTION: TISSUE CONTRACTABILITY

FILE REFERENCE: 9596-331

CURRENT APPLICATION NUMBER: US/09/880,503

CURRENT FILING DATE: 2001-06-13

PRIOR APPLICATION NUMBER: US 60/212,847

PRIOR FILING DATE: 2000-06-20

1 NUMBER OF SEQ ID NOS: 18
2 SOFTWARE: PatentIn Ver. 2.1
3 SEQ ID NO 7
4 LENGTH: 323
5 TYPE: PRT
6 ORGANISM: Homo sapiens
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

Query Match 75.5%; Score 1703; DB 10; Length 323;
Best Local Similarity 79.2%; Pred. No. 6.9e-142;
Matches 319; Conservative 2; Mismatches 2; Indels 80; Gaps 3;
QY 1 SNELHVPNSCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 1 SNELHVPNSCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 48
QY 61 KASTDTMGRPCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 120
DB 49 -----PESP-----PE-----54
QY 121 PLVQECWHD CADGK LKFCQCKTLRPRFKLIIGGEFTTIENQPMFAAIYRRHGGSVTVV 180
DB 55 -----ELKFCQCKTLRPRFKLIIGGEFTTIENQPMFAAIYRRHGGSVTVV 100
QY 181 CGGSLIPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLIHLKDYSD 240
DB 101 CGGSLIPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLIHLKDYSD 160
QY 241 TLAHNDIALKIRSKGRCAPRTIQTICLPSMYNDPOFGTSCITGFGKENS TDYLY 300
DB 161 TLAHNDIALKIRSKGRCAPRTIQTICLPSMYNDPOFGTSCITGFGKENS TDYLY 220
QY 301 PEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRM 360
DB 221 PEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRM 280
QY 361 TLGTIVSWGRGCA LKDPGVTVRSHFPLWIRSHTKENGLAL 403
DB 281 TLGTIVSWGRGCA LKDPGVTVRSHFPLWIRSHTKENGLAL 323
RESULT 12
US-10-106-698-6266
; Sequence 6266, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; CURRENT APPLICATION NUMBER: US/10/106,698
; FILE REFERENCE: PA005P1
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6266
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-6266
Query Match 65.4%; Score 1477; DB 15; Length 337;
Best Local Similarity 92.4%; Pred. No. 5.9e-122;
Matches 267; Conservative 3; Mismatches 5; Indels 14; Gaps 2;
QY 1 SNELHVPNSCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 27 SNELHVPNSCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 86

FILE REFERENCE: UTSN:022US
; CURRENT APPLICATION NUMBER: US/10/407,821
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-821-3

Query Match 64.9%; Score 1454; DB 12; Length 268;
Best Local Similarity 100.0%; Pred. No. 6.3e-121; Indels 0; Gaps 0;
Matches 268; Conservative 0; Mismatches 0;

QY 136 LKFCQCKTLRPFKIIIGGEFTTIENQPFALYRHRGGSVTVCGGSLISPCWVISAT 195
DB 1 LKFCQCKTLRPFKIIIGGEFTTIENQPFALYRHRGGSVTVCGGSLISPCWVISAT 60
QY 196 HCFIDVPKEDYIVYLGRLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRS 255
DB 61 HCFIDVPKEDYIVYLGRLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRS 120
QY 256 KEGRCQAPRTIOTICLPMSYNDPQGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHR 315
DB 121 KEGRCQAPRTIOTICLPMSYNDPQGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHR 180
QY 316 ECQPHYVSEVTTKMLCAADPOWKTDCSQGSGPLVCSLQGRMTLTGIVSWGRGCAIK 375
DB 181 ECQPHYVSEVTTKMLCAADPOWKTDCSQGSGPLVCSLQGRMTLTGIVSWGRGCAIK 240
QY 376 DKPGVTVRSHFLPWIRSHKTEENGAL 403
DB 241 DKPGVTVRSHFLPWIRSHKTEENGAL 268

RESULT 15
US-09-264-468B-2
; Sequence 2, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310.US.P1
; CURRENT APPLICATION NUMBER: US/09/264,468B
; CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-264-468B-2

Query Match 59.1%; Score 1333; DB 10; Length 246;
Best Local Similarity 99.2%; Pred. No. 2e-109;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 151 IIGGEFTTIENQPFALYRHRGGSVTVCGGSLISPCWVISATHCFIDVPKEDYIVY 210

DB 1 IIGGEFTTIENQPFALYRHRGGSVTVCGGSLISPCWVISATHCFIDVPKEDYIVY 60
QY 211 LGRSLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRSKGRCQAPRTIOTI 270
DB 61 LGRSLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRSKGRCQAPRTIOTI 120
QY 271 CLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYVSEVTTK 330
DB 121 ALPSMYNDPQGTSCETITGFGKEQSTDYLYPEQLKMTVVKLISHRECOQPHYVSEVTTK 180
QY 331 MLCAADPOWKTDCSQGSGPLVCSLQGRMTLTGIVSWGRGCAIKDKPGVTVRSHFLPW 390
DB 181 MLCAADPOWKTDCSQGSGPLVCSLQGRMTLTGIVSWGRGCAIKDKPGVTVRSHFLPW 240
QY 391 IRSHTK 396
DB 241 IRSHTK 246

RESULT 16
US-09-898-837A-47
; Sequence 47, Application US/09898837A
; Publication No. US20030077697A1
; GENERAL INFORMATION:
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Majumder, Kumud
; APPLICANT: Vernet, Corine
; APPLICANT: Herrmann, John L.
; APPLICANT: Burgess, Catherine
; APPLICANT: Fernandes, Elma
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Rastelli, Luca
; APPLICANT: CuraGen Corporation
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: MacDougall, John R.
; TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN-KINASE LIKE PROTEINS AND
; FILE REFERENCE: 15966-598 CIP
; CURRENT APPLICATION NUMBER: US/09/898,837A
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/165,986
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,839
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/195,637
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/197,080
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/181,347
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,195
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/215,906
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/715,427
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-898-837A-47

Query Match 58.4%; Score 1318; DB 11; Length 241;
Best Local Similarity 99.6%; Pred. No. 4e-108;
Matches 240; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 151 IIGGEFTTIENQPFALYRHRGGSVTVCGGSLISPCWVISATHCFIDVPKEDYIVY 210

Db 1 IIGGEFTTIENQWFAAIYRRHRRGGVTVYVCGGSLMSPCWVIGATCFIDYPKKEDIYV 60
Qy 211 LGSRLNSNTQGMKEVENLILHKDYSDTLAHNDIALLKIRSKGRCQAQPSRTIQI 270
Db 61 LGSRLNSNTQGMKEVENLILHKDYSDTLAHNDIALLKIRSKGRCQAQPSRTIQI 120
Qy 271 CLPSMYNDPQFGTSCETGFGKENSTDYLYPEOLKMTVVKLISHRECQOPHYGSEVTTK 330
Db 121 CLPSMYNDPQFGTSCETGFGKENSTDYLYPEOLKMTVVKLISHRECQOPHYGSEVTTK 180
Qy 331 MLCAADPQWTKDCQGGSGPLVCSLQGRMTLGTGIVSWGRGCALKDKPGVYTRVSHFLPW 390
Db 181 MLCAADPQWTKDCQGGSGPLVCSLQGRMTLGTGIVSWGRGCALKDKPGVYTRVSHFLPW 240
Qy 391 I 391
Db 241 I 241

RESULT 17
US-09-987-457-18
; Sequence 18, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens (tPA)
US-09-987-457-18

Query Match 38.0%; Score 858.5; DB 11; Length 527;
Best Local Similarity 37.5%; Pred. No. 3.2e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
Qy 3 ELHQP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFY 58
Db 42 QCHSVPVKSCSEPRCFNGGTCQALYFSDF-VQCPEGFAGKCEIDTRATCYEDQGISY 100
Qy 59 RGKASTDTMGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVG 118
Db 101 RGTWSTAESGAECTNWNSSALAAQPYSGRRPDAIRLGLGNHNYCRNPNDRSKPWCYVKA 160
Qy 119 LKPLVQECMVHDCADG----- 134
Db 161 GKYSSEFCSTPACSEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNP 220
Qy 135 -----KLKFO-----CG-OKTLRPRFKIIGGE 155
Db 221 AQUALGLGKHNYCRNPDGDAKWPCHVLMNRRLTWECVDPSCSTCGLRQYQOPRIKGG 280
Qy 156 FTTIENOPWFAAIYRRH-RGGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDIYVILGRS 214
Db 281 FADIASHWQAALFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPFPHLTVILGRT 340
Qy 215 RLNSNTQGMKEVENLILHKDYSDTLAHNDIALLKIRSKGRCQAQPSRTIQI 274
Db 341 YRVVPGEERQKEVEKYIVHKEFDDDT--YDNDIALQLKSDSRCAQESSVVRVTVCLPP 398

Qy 275 MYNDPQFTSCETGFGKENSTDYLYPEOLKMTVVKLISHRECQOPHYGSEVTTKMLCA 334
Db 399 ADLQLPDWTCELSGKGHEALSPFYSERLKEAHVRLYFSSRCTSOHLNLRVTVDNMLCA 458
Qy 335 AD-----PQWKT-DSCQDGGSGPLVCSLQGRMTLGTGIVSWGRGCALKDKPGVYTRVSHFL 388
Db 459 GUTRGGGQANLHDACQDGGSGPLVCLNDGRMTLVGLISWGLGCGQKDVPGVYTKVTVL 518
Qy 389 PMIRSHTK 396
Db 519 DWIRDNR 526
RESULT 18
US-09-987-455-19
; Sequence 19, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-455-19

Query Match 38.0%; Score 858.5; DB 11; Length 527;
Best Local Similarity 37.5%; Pred. No. 3.2e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;
Qy 3 ELHQP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFY 58
Db 42 QCHSVPVKSCSEPRCFNGGTCQALYFSDF-VQCPEGFAGKCEIDTRATCYEDQGISY 100
Qy 59 RGKASTDTMGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVG 118
Db 101 RGTWSTAESGAECTNWNSSALAAQPYSGRRPDAIRLGLGNHNYCRNPNDRSKPWCYVKA 160
Qy 119 LKPLVQECMVHDCADG----- 134
Db 161 GKYSSEFCSTPACSEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNP 220
Qy 135 -----KLKFO-----CG-OKTLRPRFKIIGGE 155
Db 221 AQUALGLGKHNYCRNPDGDAKWPCHVLMNRRLTWECVDPSCSTCGLRQYQOPRIKGG 280
Qy 156 FTTIENOPWFAAIYRRH-RGGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDIYVILGRS 214
Db 281 FADIASHWQAALFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPFPHLTVILGRT 340
Qy 215 RLNSNTQGMKEVENLILHKDYSDTLAHNDIALLKIRSKGRCQAQPSRTIQI 274
Db 341 YRVVPGEERQKEVEKYIVHKEFDDDT--YDNDIALQLKSDSRCAQESSVVRVTVCLPP 398
Qy 275 MYNDPQFTSCETGFGKENSTDYLYPEOLKMTVVKLISHRECQOPHYGSEVTTKMLCA 334
Db 399 ADLQLPDWTCELSGKGHEALSPFYSERLKEAHVRLYFSSRCTSOHLNLRVTVDNMLCA 458

QY 335 AD-----PQWKT-DSCQDGGGLVCSLQGRMTLTGIVSMGRGKALDKPGVYTRVSHFL 388
Db 459 GDTSGGPGQANLHDACQDGGGLVCLNDGRMTLVGLIISWGLCGGQKDPGVYTKVTNYL 518
QY 389 PWIRSHK 396
Db 519 DWIRDNMR 526

RESULT 19

US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only);
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; PRIOR FILING DATE: 2001-10-01
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-271-7

Query Match 38.0%; Score 858.5; DB 9; Length 562;
Best Local Similarity 37.5%; Pred. No. 3.4e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYPSNIHWNCPCPKFGGQCHCEIDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSEPRCFNGGTCQALYFSDF-VQCPEGFAGKCEIDTRATCYEDQGISY 135

QY 59 RGAASDTMTGRPCLPNWSATVLQOTYHAHRSDALQGLGKHNYCRNPDNRPRPCVYQVG 118
Db 136 RGTWTAESGAECTNWNSSALAQKPYSGRRPDARLGLGNHNYCRNPDNRDRSKPCYVFXA 195

QY 119 LKPLVQECMVHDCADG-----KLKFO-----CG-QKTLRPRFKIIGGE 155

Db 196 GKYSSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLRWNSMILIGKYVTAQNP 255

QY 135 -----KLKFO-----CG-QKTLRPRFKIIGGE 155

Db 256 AQAALGLKHNYCRNPDGDAKPMCHVKNRRLTWECYDVPSCSTCGLRQYSQPFRIKGG 315

QY 156 FTTIENQPFWFAIYRRH-RGGSVTVVCGSLISPCWVISATHCFIDYPPKEDYIVVLGRS 214

Db 316 FADIASHPWQAALFAKHRRSPGERFLCGGILISSCWILSAHCFQERPPHLLTVILGR 375

QY 215 RLNSNTQGENKFEVENLILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIQTICLPS 274

Db 376 YRVVPGEEQKFEVEKYIVHKEFDDDT--YNDIALQLKSDSSRCAQESSVVRTVCLPP 433

QY 275 MYNDPQGTSCETGFCENSTDYLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLCA 334

Db 434 ADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYPSRRTSQHLLNRTVTDNMLCA 493

QY 335 AD-----PQWKT-DSCQDGGGLVCSLQGRMTLTGIVSMGRGKALDKPGVYTRVSHFL 388

Db 494 GDTSGGPGQANLHDACQDGGGLVCLNDGRMTLVGLIISWGLCGGQKDPGVYTKVTNYL 553

QY 389 PWIRSHK 396

Db 554 DWIRDNMR 561

RESULT 20

US-09-974-298-145

; Sequence 145, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 145
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CD1
US-09-974-298-145

Query Match 38.0%; Score 858.5; DB 10; Length 562;
Best Local Similarity 37.5%; Pred. No. 3.4e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYPSNIHWNCPCPKFGGQCHCEIDKSKTCYEGNGHFY 58

Db 77 QCHSVPVKSCSEPRCFNGGTCQALYFSDF-VQCPEGFAGKCEIDTRATCYEDQGISY 135

QY 59 RGAASDTMTGRPCLPNWSATVLQOTYHAHRSDALQGLGKHNYCRNPDNRPRPCVYQVG 118

Db 136 RGTWTAESGAECTNWNSSALAQKPYSGRRPDARLGLGNHNYCRNPDNRDRSKPCYVFXA 195

QY 119 LKPLVQECMVHDCADG-----KLKFO-----CG-QKTLRPRFKIIGGE 155

Db 196 GKYSSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLRWNSMILIGKYVTAQNP 255

QY 135 -----KLKFO-----CG-QKTLRPRFKIIGGE 155

Db 256 AQAALGLKHNYCRNPDGDAKPMCHVKNRRLTWECYDVPSCSTCGLRQYSQPFRIKGG 315

QY 156 FTTIENQPFWFAIYRRH-RGGSVTVVCGSLISPCWVISATHCFIDYPPKEDYIVVLGRS 214

Db 316 FADIASHPWQAALFAKHRRSPGERFLCGGILISSCWILSAHCFQERPPHLLTVILGR 375

QY 215 RLNSNTQGENKFEVENLILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIQTICLPS 274

Db 376 YRVVPGEEQKFEVEKYIVHKEFDDDT--YNDIALQLKSDSSRCAQESSVVRTVCLPP 433

QY 275 MYNDPQGTSCETGFCENSTDYLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLCA 334

Db 434 ADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYPSRRTSQHLLNRTVTDNMLCA 493

QY 335 AD-----PQWKT-DSCQDGGGLVCSLQGRMTLTGIVSMGRGKALDKPGVYTRVSHFL 388

Db 494 GDTSGGPGQANLHDACQDGGGLVCLNDGRMTLVGLIISWGLCGGQKDPGVYTKVTNYL 553

QY 389 PWIRSHK 396

Db 554 DWIRDNMR 561

RESULT 21

US-10-443-701-4

; Sequence 4, Application US/10443701
; Publication No. US20030199016A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Yuan
; TITLE OF INVENTION: REVERSE-PHASE HPLC ASSAY FOR PLASMINOGEN ACTIVATORS
; FILE REFERENCE: P1788R1
; CURRENT APPLICATION NUMBER: US/10/443,701
; CURRENT FILING DATE: 2003-05-21
; PRIOR APPLICATION NUMBER: US/09/703,695

```
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/163,607
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-443-701-4

Query Match      38.0%; Score 858.5; DB 12; Length 562;
Best Local Similarity 37.5%; Pred. No. 3.4e-67;
Matches 183; Conservative 56; Mismatches 152; Indels 97; Gaps 10;

QY 3 ELHQP-SNCD---CLNGGTCVSNKYFNIHWCNPKKFGQHQCEIDSKSKTCYEGNGHFY 58
Db 77 QCHSVPKSCSEPCFNGGTCQALYFSDP-VQCQPEGAGKCEIDTRATCYEDQGISY 135
QY 59 RGKASTDTMGRCLPWSNATVLOQTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVG 118
Db 136 RGTWSTAESGAECTNWNSSALAQPKYSGRRPDAILRLGNHNYCRNPNDRDCKPWCYVFK 195
QY 119 LKPLVQECMVHDCADG----- 134
Db 196 KYSSSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNP 255
QY 135 -----KLKFO-----CG-QKTLRPRFKIIGGE 155
Db 256 AQALGLGKHNYCRNPDGDAKPMCHVKNRRLTWECYDVPSCSTCGLRQYQPOFRIKGG 315
QY 156 FTTIENQWFAAIYRRH-RGGSVTVYCGGSLISPCWISATHCFIDYPKKEDYIVVLGR 214
Db 316 FADIASHPWQAIAFAKRRSPGERFLCGILISSCWILSAHCFQERPPPHLTIVILGR 375
QY 215 RLNSNTQGMKFEVENILHKDYSADTLAHNDIALLKIRSKRCAQPSRTIQTICLPS 274
Db 376 YRVVPGEEQKFEVEKYVHKEFDDT--YNDIALQLKSDSRCAQESSVVRTVCLPP 433
QY 275 MYNDPQGTSCITGFGKENSTDLYPEQLKMTVVKLISHRECQOPHYVSEVTTMCLCA 334
Db 434 ADLQPDWTECELSGKGHEALSPFYSERLKEAHVRLYPSRCSQHLNRTVTDNMLCA 493
QY 335 AD-----PWKLT-DSCQDSDGGPLVCSLQGRMTLTGIVSWGRCALCKPQVYTRVSHFL 388
Db 494 GDTSSGGPQANLHDACQDSDGGPLVCLNDGRMTLVGIISWGLGCGQKDPGVYTKVTNYL 553
QY 389 PWIRSHTK 396
Db 554 DWIRDNR 561

RESULT 23
US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-4

Query Match      35.1%; Score 793; DB 10; Length 135;
Best Local Similarity 100.0%; Pred. No. 3.6e-62;
```


Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDLNGTGVSNKYFNSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60
Db 1 SNELHQPNSCDLNGTGVSNKYFNSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120
Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
Db 121 PLVQECMVHDCADGK 135

RESULT 24

US-09-984-186-12
Sequence 12, Application US/09984186
Patent No. US2002015101A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: S792006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-09-984-186-12

Query Match 35.1%; Score 793; DB 10; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.6e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDLNGTGVSNKYFNSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60
Db 4 SNELHQPNSCDLNGTGVSNKYFNSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120
Db 64 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPMCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 25

US-10-237-667-12
Sequence 12, Application US/10237667
Publication No. US20030022308A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,667
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-10-237-667-12

Query Match 35.1%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.6e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 4 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPONRRPWCYVQVGLK 120
Db 64 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPONRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 26
US-10-237-708-12
; Sequence 12, Application US/10237708
; Publication No. US20030036170A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcoia Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,708
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-708-12

Query Match 35.1%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.6e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60

Db 4 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPONRRPWCYVQVGLK 120
Db 64 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPONRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 27
US-10-237-866-12
; Sequence 12, Application US/10237866
; Publication No. US20030036171A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcoia Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,866
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12

Query Match 35.1%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.6e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60

Db 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 63
Qy 61 KASDTMTGRCPLPNSATVLQOQTYHAHRSDALQGLGKHNCRPNDRRPPWCYVOVGLK 120
Db 64 KASDTMTGRCPLPNSATVLQOQTYHAHRSDALQGLGKHNCRPNDRRPPWCYVOVGLK 123
Qy 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 28

US-10-237-871-12
; Sequence 12, Application US/10237871
; Publication No. US20030036172A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,871
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-10-237-871-12

Query Match 35.1%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.6e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
Db 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 63

Qy 61 KASDTMTGRCPLPNSATVLQOQTYHAHRSDALQGLGKHNCRPNDRRPPWCYVOVGLK 120
Db 64 KASDTMTGRCPLPNSATVLQOQTYHAHRSDALQGLGKHNCRPNDRRPPWCYVOVGLK 123
Qy 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 29

US-10-237-624-12
; Sequence 12, Application US/10237624
; Publication No. US20030082747A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,624
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-10-237-624-12

Query Match 35.1%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.6e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
Db 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 63

QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
Db |||||
64 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
Db |||||
124 PLVQECMVHDCADGK 138

RESULT 30
US-09-880-503-8
; Sequence 8, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-8

Query Match 35.1%; Score 793; DB 10; Length 143;
Best Local Similarity 100.0%; Pred. No. 3.8e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60
Db |||||
1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
Db |||||
61 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
Db |||||
121 PLVQECMVHDCADGK 135

Search completed: December 3, 2003, 15:05:45
Job time : 42.3738 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:35:43 ; Search time 7.39655 Seconds
(without alignments)
818.010 Million cell updates/sec

Title: US-09-880-503-8

Perfect score: 837

Sequence: 1 SNEHQVPSNCDCLNGTCTV.....QECVHDCADGKKPSPPEE 143

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	837	100.0	200	4	US-09-101-272G-73
2	837	100.0	208	4	US-09-101-272G-98
3	837	100.0	411	1	US-08-087-163-1
4	837	100.0	411	1	US-08-286-748B-18
5	837	100.0	411	1	US-08-153-799-18
6	837	100.0	430	1	US-07-942-157A-3
7	837	100.0	431	4	US-09-101-272G-1
8	837	100.0	431	6	518829-1
9	827	98.8	411	3	US-09-181-816-1
10	826	98.7	411	2	US-08-560-098A-48
11	821.5	98.1	430	6	5219569-2
12	813	97.1	157	3	US-08-142-590B-25
13	793	94.7	138	2	US-08-797-689-12
14	788	94.1	194	4	US-09-101-272G-80
15	788	94.1	201	4	US-09-101-272G-96
16	559	66.8	432	2	US-08-560-098A-47
17	558	66.7	365	1	US-08-093-741-83
18	558	66.7	365	2	US-08-720-012-83
19	558	66.7	393	1	US-08-560-098A-44
20	558	66.7	393	3	US-08-967-024C-24
21	558	66.7	393	3	US-08-967-024C-25
22	513	61.3	89	4	US-09-101-272G-62
23	335.5	40.1	477	2	US-08-560-098A-51
24	328.5	39.2	527	1	US-07-609-510B-16
25	328.5	39.2	527	2	US-08-811-949-39
26	328.5	39.2	527	5	PCT-US91-01025A-2
27	328.5	39.2	527	6	5185259-8

28	328.5	39.2	527	6	5520913-1	Patent No. 5520913
29	328.5	39.2	546	6	5200340-6	Patent No. 5200340
30	328.5	39.2	562	2	US-08-811-949-43	Sequence 43, Appl
31	328.5	39.2	562	2	US-08-560-098A-50	Sequence 50, Appl
32	328.5	39.2	562	2	US-08-883-795A-38	Sequence 38, Appl
33	328.5	39.2	562	6	5185259-3	Patent No. 5185259
34	328.5	39.2	562	6	5200340-2	Patent No. 5200340
35	328.5	39.2	562	6	5344773-2	Patent No. 5344773
36	321.5	38.4	562	6	5244676-5	Patent No. 5244676
37	300	35.8	233	3	US-08-438-745-15	Sequence 15, Appl
38	300	35.8	233	3	US-08-438-745-17	Sequence 17, Appl
39	300	35.8	233	3	US-09-219-019-15	Sequence 15, Appl
40	300	35.8	233	3	US-09-219-019-17	Sequence 17, Appl
41	300	35.8	233	5	PCT-US94-05669A-15	Sequence 15, Appl
42	300	35.8	233	5	PCT-US94-05669A-17	Sequence 17, Appl
43	300	35.8	233	3	US-08-438-745-13	Sequence 13, Appl
44	300	35.8	235	3	US-09-219-019-13	Sequence 13, Appl
45	300	35.8	235	5	PCT-US94-05669A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-101-272G-73
; Sequence 73, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: ATP domain of uPA
US-09-101-272G-73

Query Match	100.0%	Score 837;	DB 4;	Length 200;
Best Local Similarity	100.0%	Pred. No. 4.2e-76;		
Matches 143;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	SNEHQVPSNCDCLNGTCTVSNKYFSNIHWCNPKYFGQHCEIDKSKTCYEGNGHFYRG	60	
Db	21	SNEHQVPSNCDCLNGTCTVSNKYFSNIHWCNPKYFGQHCEIDKSKTCYEGNGHFYRG	80	
Qy	61	KASTDTMGPRCPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVOVGLK	120	
Db	81	KASTDTMGPRCPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVOVGLK	140	
Qy	121	PLVQECMVHDCADGKKPSPPEE	143	
Db	141	PLVQECMVHDCADGKKPSPPEE	163	
RESULT 2				
US-09-101-272G-98				
; Sequence 98, Application US/09101272G				
; Patent No. 6509445				
; GENERAL INFORMATION:				
; APPLICANT: Nissin Food Products Co., Ltd.				
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR				
; FILE REFERENCE: Q50979				
; CURRENT APPLICATION NUMBER: US/09/101,272G				
; CURRENT FILING DATE: 1998-07-08				

```

; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 98
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATH1-ML chimeric protein
US-09-101-272G-98

Query Match      100.0%; Score 837; DB 4; Length 208;
Best Local Similarity 100.0%; Pred. No. 4.4e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
Db 2 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 61
Qy 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPWCYVQVGLK 120
Db 62 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPWCYVQVGLK 121
Qy 121 PLVQECMVHDCADGKKPSPPEE 143
Db 122 PLVQECMVHDCADGKKPSPPEE 144

RESULT 3
US-08-087-163-1
; Sequence 1, Application US/08087163
; Patent No. 5472692
; GENERAL INFORMATION:
; APPLICANT: Liu, Jian-Ning
; APPLICANT: Gurewicz, Victor
; TITLE OF INVENTION: PRO-UKINASE MUTANTS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087.163
; FILING DATE: 07/02/93
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04353/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
US-08-087-163-1

Query Match      100.0%; Score 837; DB 4; Length 208;
Best Local Similarity 100.0%; Pred. No. 4.4e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
Db 2 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 61
Qy 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPWCYVQVGLK 120
Db 62 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPWCYVQVGLK 121
Qy 121 PLVQECMVHDCADGKKPSPPEE 143
Db 122 PLVQECMVHDCADGKKPSPPEE 144

RESULT 4
US-08-286-748B-18
; Sequence 18, Application US/08286748B
; Patent No. 5759542
; GENERAL INFORMATION:
; APPLICANT: Victor Gurewicz
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
; OF DRUGS BY PLATELETS FOR THE TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,748B
; FILING DATE: August 5, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: J. Peter Fasse
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04547/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-286-748B-18

Query Match      100.0%; Score 837; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 9.4e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
Db 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
Qy 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPWCYVQVGLK 120
Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPWCYVQVGLK 120
```

QY 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 121 PLVQECMVHDCADGKKPSSPPEE 143

RESULT 5
US-08-153-799-18
; Sequence 18, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-153-799-18

Query Match 100.0%; Score 837; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 9.4e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKXFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKXFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWSATVLQOQTYHAHRSALQLGLGKHNCRNPNRRRPMWCYVQVGLK 120
Db 61 KASTDTMGRPCLPWSATVLQOQTYHAHRSALQLGLGKHNCRNPNRRRPMWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSSPPEE 143

Db 121 PLVQECMVHDCADGKKPSSPPEE 143

RESULT 6
US-07-942-157A-3
; Sequence 3, Application US/07942157A
; Patent No. 5648253
; GENERAL INFORMATION:
; APPLICANT: Wei, Cha-Mer
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,157A
; FILING DATE: 19920908
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631673
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: TSI108Cont.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 198..203
; OTHER INFORMATION: /label= modified
; OTHER INFORMATION: /note= "six amino acids deleted in mutant"
US-07-942-157A-3

Query Match 100.0%; Score 837; DB 1; Length 430;
Best Local Similarity 100.0%; Pred. No. 9.9e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKXFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 20 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKXFGGQHCIEDKSKTCYEGNGHFYRG 79
QY 61 KASTDTMGRPCLPWSATVLQOQTYHAHRSALQLGLGKHNCRNPNRRRPMWCYVQVGLK 120
Db 80 KASTDTMGRPCLPWSATVLQOQTYHAHRSALQLGLGKHNCRNPNRRRPMWCYVQVGLK 139
QY 121 PLVQECMVHDCADGKKPSSPPEE 143
Db 140 PLVQECMVHDCADGKKPSSPPEE 162

'SULT 7
 '09-101-272G-1
 Sequence 1, Application US/09101272G
 Patent No. 6509445
 GENERAL INFORMATION:
 APPLICANT: Nissin Food Products Co., Ltd.
 TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
 FILE REFERENCE: Q50979
 CURRENT APPLICATION NUMBER: US/09/101.272G
 CURRENT FILING DATE: 1998-07-08
 PRIOR APPLICATION NUMBER: JP 1059/1996
 PRIOR FILING DATE: 1996-01-08
 NUMBER OF SEQ ID NOS: 107
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 1
 LENGTH: 431
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: mat peptide
 LOCATION: (21)..
 OTHER INFORMATION:
 NAME/KEY: misc:feature
 LOCATION: (20)..
 OTHER INFORMATION: Urokinase-type plasminogen act
 '09-101-272G-1

```

RESULT 9
US-09-181-816-1
; Sequence 1, Application US/09181816
; Patent No. 6277818
; GENERAL INFORMATION:
; APPLICANT: MAZAR, Andrew P.
; APPLICANT: JONES, Terence R.
; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
; TITLE OF INVENTION: PLASMINOGEN ACTIVATOR RECEPTOR
; FILE REFERENCE: 329042000300 SIDN 1-7
; CURRENT APPLICATION NUMBER: US/09/181,816
; CURRENT FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-181-816-1

```

RESULT 10
US-08-560-098A-48
Sequence 48, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WNEINDT, Stephan
APPLICANT: WEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins Having Fibrinolytic and
Coagulation-Inhibiting Properties
TITLE OF INVENTION: 60
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,98A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800


```
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-48

Query Match
Best Local Similarity 98.7%; Score 826; DB 2; Length 411;
Matches 142; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWNASATVLOQTYHAHRSDALQGLGKHNYCRNPDNRPRPCVYQVGLK 120
DB 61 KASTDTMGPRCLPWNASATVLOQTYHAHRSDALQGLGKHNYCRNPDNRPRPCVYQVGLK 120
QY 121 PLVQECMVHDCADGKKPSSPPEE 143
DB 121 PLVQECMVHDCADGKKPSSPPEE 143

RESULT 11
5219569-2
; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO: 2:
; LENGTH: 430
5219569-2

Query Match
Best Local Similarity 98.1%; Score 821.5; DB 6; Length 430;
Matches 142; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGPRCLPWNASATVLOQTYHAHRSDALQGLGKHNYCRNPDNRPRPCVYQVGLK 120
DB 81 KASTDTMGPRCLPWNASATVLOQTYHAHRSDALQGLGKHNYCRNPDNRPRPCVYQVGLK 140
QY 121 PLVQECMVHDCADGKKPSSPPEE 143
DB 141 PLVQECMVHDCADG-KPSSPPEE 162

RESULT 12
US-08-142-590B-25
; Sequence 25, Application US/08142590B
; Patent No. 6120765
; GENERAL INFORMATION:
; APPLICANT: HIBINO, Tashihiko; TAKAHASHI, Tadabito; HORII, Izumi; and
; GOETINCK,
; TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
```

```
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,590B
; FILING DATE: 25-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,318
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: MGP-009CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 157 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-142-590B-25

Query Match
Best Local Similarity 97.1%; Score 813; DB 3; Length 157;
Matches 139; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWNASATVLOQTYHAHRSDALQGLGKHNYCRNPDNRPRPCVYQVGLK 120
DB 61 KASTDTMGPRCLPWNASATVLOQTYHAHRSDALQGLGKHNYCRNPDNRPRPCVYQVGLK 120
QY 121 PLVQECMVHDCADGKKPSSPPEE 143
DB 121 PLVQECMVHDCADGKKPSSPPEE 143

RESULT 13
US-08-797-689-12
; Sequence 12, Application US/08797689
; Patent No. 5878969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-797-689-12

Query Match          94.7%; Score 793; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 6.7e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKKGGOHCEIDKSKTCYEGNGHYRG 60
Db 4 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKKGGOHCEIDKSKTCYEGNGHYRG 63

Qy 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCWYVQVGLK 120
Db 64 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCWYVQVGLK 123

Qy 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 14
US-09-101-272G-80
; Sequence 80; Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101.272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 80
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATFI chimeric protein
; US-09-101-272G-80

Query Match          94.1%; Score 788; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 3.1e-71;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKKGGOHCEIDKSKTCYEGNGHYRG 60
Db 2 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKKGGOHCEIDKSKTCYEGNGHYRG 61

Qy 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCWYVQVGLK 120
Db 62 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCWYVQVGLK 121

Qy 121 PLVQECMVHDCADGK 135
Db 122 PLVQECMVHDCADGK 135

RESULT 15
US-09-101-272G-96
; Sequence 96; Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101.272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 96
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATFI-CL chimeric protein
; US-09-101-272G-96

Query Match          94.1%; Score 788; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 3.2e-71;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKKGGOHCEIDKSKTCYEGNGHYRG 60
Db 2 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKKGGOHCEIDKSKTCYEGNGHYRG 61

Qy 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCWYVQVGLK 120
Db 62 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCWYVQVGLK 121

Qy 121 PLVQECMVHDCADGK 134
Db 122 PLVQECMVHDCADGK 135

RESULT 16
US-08-560-098A-47
; Sequence 47; Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENDET, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; US-08-560-098A-47
```

```

; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, JOSEPH D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-47

Query Match 66.8%; Score 559; DB 2; Length 432;
Best Local Similarity 77.1%; Pred. No. 5.5e-48;
Matches 101; Conservative 8; Mismatches 16; Indels 5; Gaps 1;

QY 13 CLNGGTCVSNKYFVSIHNCNPKFKGGHCEIDKSKTCYEGNGHYRGKASTDTMGRPCL 72
DB 40 CVTGGTPEKSHNGDFEIEPEEY-----LQISKTCYEGNGHYRGKASTDTMGRPCL 93
QY 73 PMSATVLTQOTYHARSALQGLGKHNYCRPNRRPWCYVQVGLKPLVQECMVHCA 132
DB 94 PMSATVLTQOTYHARSALQGLGKHNYCRPNRRPWCYVQVGLKPLVQECMVHCA 153
QY 133 DGKXPSSPPEE 143
DB 154 DGKXPSSPPEE 164

RESULT 17
US-08-093-741-83
; Sequence 83, Application US/08093741
; Patent No. 5681721
; GENERAL INFORMATION:
; APPLICANT: STEFFENS, GERD J.
; APPLICANT: WENDT, STEPHAN
; APPLICANT: SCHNEIDER, JOHANNES
; APPLICANT: HEINZEL-WIELAND, REGINA
; APPLICANT: SAUNDERS, DEREK J.
; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N. W. Suite 700
; CITY: Washington, D.C.
; COUNTRY: U.S.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/093,741
; FILING DATE: 20-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/093,741
; FILING DATE: 20-JUL-1993
; APPLICATION NUMBER: DE P43 23 754.1
; FILING DATE: 15-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, JOSEPH D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/41345
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)628-8800
; TELEFAX: (202)628-8844
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-093-741-83

```

```

; REFERENCE/DOCKET NUMBER: 148/41345
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)628-8800
; TELEFAX: (202)628-8844
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-093-741-83

Query Match 66.7%; Score 558; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 5.7e-48;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHYRGKASTDTMGRPCLPWSATVLTQOTYHARSALQGLGKHNYCRNPD 106
DB 1 SKTCYEGNGHYRGKASTDTMGRPCLPWSATVLTQOTYHARSALQGLGKHNYCRNPD 60
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKXPSSPPEE 143
DB 61 NRRRPWCYVQVGLKPLVQECMVHDCADGKXPSSPPEE 97

RESULT 18
US-08-720-012-83
; Sequence 83, Application US/08720012
; Patent No. 5747291
; GENERAL INFORMATION:
; APPLICANT: STEFFENS, GERD J.
; APPLICANT: WENDT, STEPHAN
; APPLICANT: SCHNEIDER, JOHANNES
; APPLICANT: HEINZEL-WIELAND, REGINA
; APPLICANT: SAUNDERS, DEREK J.
; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N. W. Suite 700
; CITY: Washington, D.C.
; COUNTRY: U.S.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,012
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/093,741
; FILING DATE: 20-JUL-1993
; APPLICATION NUMBER: DE P43 23 754.1
; FILING DATE: 15-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, JOSEPH D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/41345
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)628-8800
; TELEFAX: (202)628-8844
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-720-012-83

Query Match 66.7%; Score 558; DB 1; Length 365;

```

Best Local Similarity 100.0%; Pred. No. 5.7e-48; Indels 0; Gaps 0;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPD 106
Db 1 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPD 60

QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKSPSPPEE 143
Db 61 NRRRPWCYVQVGLKPLVQECMVHDCADGKKSPSPPEE 97

RESULT 19
US-08-560-098A-44
; Sequence 44, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: JANOSCH, Elke
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560.098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-44
Query Match 66.7%; Score 558; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 6.2e-48;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPD 106
Db 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPD 61

QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKSPSPPEE 143
Db 62 NRRRPWCYVQVGLKPLVQECMVHDCADGKKSPSPPEE 98

RESULT 20
US-08-967-024C-24
; Sequence 24, Application US/08967024C

Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSCH, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967.024C
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 42 665.8
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42444
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-967-024C-24
Query Match 66.7%; Score 558; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 6.2e-48; Indels 0; Gaps 0;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPD 106
Db 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPD 61

QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKSPSPPEE 143
Db 62 NRRRPWCYVQVGLKPLVQECMVHDCADGKKSPSPPEE 98

RESULT 21
US-08-967-024C-25
; Sequence 25, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSCH, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 66.7%; Score 558; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 6.2e-48;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNP 106
DB 2 SKTCYEGNHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNP 61

QY 107 NRRPRPCYYQVGLKPLVQECMVHDCADGKKPSPPEE 143
DB 62 NRRPRPCYYQVGLKPLVQECMVHDCADGKKPSPPEE 98

RESULT 22
US-09-101-272G-62
; Sequence 62, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; FILE REFERENCE: Q50979
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 62
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: residues 43-131 of the ATF domain of UPA
US-09-101-272G-62

Query Match 61.3%; Score 513; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 3.5e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 EIDKSKTCYEGNHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNP 102
DB 1 EIDKSKTCYEGNHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNP 60

QY 103 RNPENRRPCYYQVGLKPLVQECMVHDC 131

DB 61 RNPENRRPCYYQVGLKPLVQECMVHDC 89

RESULT 23
US-08-560-098A-51
; Sequence 51, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNELOT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-51

Query Match 40.1%; Score 335.5; DB 2; Length 477;
Best Local Similarity 46.3%; Pred. No. 1.2e-25;
Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;

QY 3 ELHQVP-----SNCDCLNGGTQVSNKYFSNIHWCNPKXFGGQHCEIDKSKTCYEGNHFY 58
DB 78 QCHTFPVKSCSELRCFNGGTCWQAASFDF-VCQPKGYTGKQCEVDTHATCYKDDQGVY 136

QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRNPENRRPCYYQVG 118
DB 137 RGTWSTSESGAOCINWNSNLLTRTYNGRRSDAITLGLGNHNYCRNPENRRPCYYVKA 196

QY 119 LKPLVQECMVHDC 132
DB 197 SKFILEFCSPVCS 210

RESULT 24
US-07-609-510B-16
; Sequence 16, Application US/07609510B
; Patent No. 5326700
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue
; NUMBER OF SEQUENCES: 16

Query Match	39.2%	Score 328.5	DB 2	Length 527
Best Local Similarity	46.3%	Pred.No. 6.9e-25		
Matches 63	Conservative 14	Mismatches 5	Indels 5	Gaps 3

QY	3	ELHQPV-SNCD---	CLANGGTCVSNKYFSNIHWNCNCPKKGQGHCEIDKSKTCYENGHPY	58
Db	42	QCHSVFVKSCEPRFCNGGTCQQAQYEDF-VQCPBEGFAGKCCETDTRATCYEQGISY	100	
QY	59	RKASGTDIMGRPLPWSNATVLOQTYHAHRSDALQLGLGKHNYCRNPONRRPWCVVQVG	118	
Db	101	RTWCTATSCGFCNNWSSALAKPVCSCPEDAIRICGLNHNVCNRPEDRSPKCVFKA	160	

```

QY      119 LKPLVOECMVHDCADG 134      : :
      : : : : : : : : : : : :
Db      161 GKYSSFCSTPACSEG 176

```

RESULT 26
 PCT-US91-01025A-2
 ; Sequence 2, Application PC/TUS9101025A
 ; GENERAL INFORMATION:
 ; APPLICANT: Genentech, Inc.
 ; TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
 ; TITLE OF INVENTION: Specific Properties
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080

CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/01025A
FILING DATE: 19910214
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/486,657
FILING DATE: 1 March 1990
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 454P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/266-1896
TELEFAX: 415/952-9891
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US91-01025A-2

CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 562 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-43

Query Match 39.2%; Score 328.5; DB 2; Length 562;
Best Local Similarity 46.3%; Pred. No. 7.4e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;
QY 3 ELHQVP-SNCD---CLNGTCVSNKYFSNIHWCNCPKKGCGHCEIDKSKTCYEGNGHFY 58
DB 77 QCHSVPVKSCSEPRCFNGGTCQQAIFYSDF-VQCPEGFAGKCEIDTRATCYEDQGISY 135
QY 59 RKASTDITNGRCLPWNSATVLQOTYHAHRSDALQLGLGKHYCNRPDNRPRPCWYVQVG 118
DB 136 RGTWSTAESGAECTWNNSALAKPKYSGRRPDALRLGLGNHNYCRNPDRDSKPCWYVFK 195
QY 119 LKPLVQECMVHDCADG 134
DB 196 GKYSEFCSTPACSEG 211

Search completed: December 3, 2003, 14:45:37
Job time : 7.39655 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:23 ; Search time 17.5948 Seconds
(without alignments)
2097.294 Million cell updates/sec

Title: US-09-880-503-8

Perfect score: 837

Sequence: 1 SNELHQPVSNCCLNGTGV.....QECMVHDCADGKKPSPPEE 143

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_protist:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	778	93.0	154	4 Q96SE8	Q96se8 homo sapien
2	627	74.9	433	6 Q8WIL0	Q8wilo cryptolagus
3	627	74.9	433	6 Q8WHY7	Q8wh7 cryptolagus
4	576	68.8	157	6 Q9TVA8	Q9tva8 bos taurus
5	569	68.0	231	11 Q8C6L2	Q8c6l2 mus musculus
6	337.5	40.3	214	6 Q8C6L2	Q8c6l2 mus musculus
7	322.5	38.5	516	4 Q8BU99	Q8bu99 homo sapien
8	315.5	37.7	559	11 Q91VP2	Q91vp2 mus musculus
9	311.5	37.2	562	6 Q8SQ23	Q8sq23 sus scrofa
10	301.5	36.0	564	6 Q8MK81	Q8mk81 cryptolagus
11	285	31.7	517	11 Q8K0B2	Q8kd02 mus musculus
12	263.5	31.5	560	4 Q4520	Q4520 homo sapien
13	258	30.8	616	6 Q97507	Q97507 sus scrofa
14	250.5	29.9	653	11 Q8VCS4	Q8vcs4 mus musculus
15	250	29.9	597	11 Q35727	Q35727 mus musculus
16	233	27.8	395	4 Q9BZW1	Q9bzw1 homo sapien

17	233	27.8	615	4 Q81Z25	Q81zz5 homo sapien
18	218.5	26.1	90	4 Q8NG20	Q8ng20 homo sapien
19	197.5	23.6	202	13 Q90675	Q90675 gallus gall
20	174	20.8	421	13 Q8AXX3	Q8axx3 xenopus lae
21	174	20.8	947	13 Q8AXY6	Q8axy6 gallus gall
22	166	19.8	716	13 Q91691	Q91691 xenopus lae
23	165.5	19.8	359	6 Q8WMR1	Q8wmr1 canis famil
24	165.5	19.8	812	11 Q9RCW3	Q9rcw3 rattus norv
25	159	19.0	452	13 Q90Y90	Q90y90 xenopus lae
26	158.5	18.9	333	4 Q8BR86	Q8br86 homo sapien
27	158.5	18.9	937	11 Q8BNP9	Q8bnp9 mus musculu
28	158.5	18.9	937	11 Q8BG10	Q8bg10 mus musculu
29	157	18.9	454	6 Q46506	Q46506 papio hamad
30	157	18.8	113	4 Q9UIR5	Q9uir5 homo sapien
31	156.5	18.7	810	4 Q15146	Q15146 homo sapien
32	155	18.5	812	11 Q91WJ5	Q91wj5 mus musculu
33	154	18.4	801	11 Q8XQ8	Q8xq8 mus musculu
34	154	18.4	806	6 Q18783	Q18783 macropus eu
35	154	18.4	944	11 Q8C3W2	Q8c3w2 mus musculu
36	154	18.4	944	11 Q8BSP6	Q8bsp6 mus musculu
37	153	18.3	113	4 Q9UIR7	Q9uir7 homo sapien
38	152.5	18.2	717	13 P70006	P70006 xenopus lae
39	152	18.2	567	4 Q13208	Q13208 homo sapien
40	152	18.2	648	4 Q9HIV4	Q9hiv4 homo sapien
41	150.5	18.0	930	13 Q8AV69	Q8av69 xenopus lae
42	148.5	17.7	381	4 Q8N2J4	Q8n2j4 homo sapien
43	148.5	17.7	399	4 Q86GL8	Q86gl8 homo sapien
44	148.5	17.7	420	4 Q9BTP9	Q9btp9 homo sapien
45	148.5	17.7	424	4 Q8NCW1	Q8ncw1 homo sapien

ALIGNMENTS

RESULT 1

Q96SE8

ID Q96SE8 PRELIMINARY; PRT; 154 AA.

AC Q96SE8; TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Urokinase-type plasminogen activator amino-terminal fragment.

GN AUF.

OS Homo sapiens (Human).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Fu J., Bai X., Ruan C.;

RT "Cloning and expression of the amino-terminal fragment of human urokinase-type plasminogen activator.",

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Bai X., Fu J., Wang W., Xi X., Ruan C.;

RT "Overexpression of the amino-terminal fragment of human urokinase-type plasminogen activator in breast cancer cells results in decreased tumor invasion, growth and angiogenesis".

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; AY029537; AAK38734.1; -.

DR InterPro; IPR006209; EGF-like.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR000001; Kringle.

DR Pfam; PF00051; Kringle; 1.

DR PRINTS; PR00016; KRINGLE.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00181; EGF; 1.

DR SMART; SM00130; KR; 1.

DR PROSITE; PS00022; EGF; 1.

DR PROSITE; PS00021; KRINGLE; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

KW Glycoprotein; Kinase; Kringle.
SQ SEQUENCE 154 AA; 17305 MW; A3CF2FCFF505572 CRC64;
Query Match 93.0%; Score 778; DB 4; Length 154;
Best Local Similarity 99.3%; Pred. No. 2.2e-76;
Matches 133; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 SNELHGVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHYRG 60
Db 21 SNELHGVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHYRG 80
Qy 61 KASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQVGLK 140
Qy 121 PLVQECMVHDCADG 134
Db 141 LLVQECMVHDCADG 154
RESULT 2
Q8MHILO PRELIMINARY; PRT; 433 AA.
ID Q8MHILO
AC Q8MHILO
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN PLAU.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22155945; PubMed=12149463;
RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
RA Dicke D.A.;
RT "Increased expression of urokinase during atherosclerotic lesion
development causes arterial constriction and lumen loss, and
accelerates lesion growth."
RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY122285; AA83187.1; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65E64F36415549B0 CRC64;
Query Match 74.9%; Score 627; DB 6; Length 433;
Best Local Similarity 76.2%; Pred. No. 1.5e-59;
Matches 109; Conservative 11; Mismatches 21; Indels 2; Gaps 1;
Qy 1 SNELHGV--PSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHY 58
Db 21 SNELHGVSDASNGCCLNGGTCVTKYFSNIWRNCNCPKFGQHCCEIDTLKTCYHGDGHSY 80
Qy 59 RGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQV 118
Db 81 RGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQV 140
Qy 119 LKPLVQECMVHDCADGKKPSSPP 141
Db 141 LKQLIQECKVHDCSSGKKPALPP 163

KW Glycoprotein; Kinase; Kringle.
SQ SEQUENCE 154 AA; 17305 MW; A3CF2FCFF505572 CRC64;
Query Match 93.0%; Score 778; DB 4; Length 154;
Best Local Similarity 99.3%; Pred. No. 2.2e-76;
Matches 133; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 SNELHGVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHYRG 60
Db 21 SNELHGVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHYRG 80
Qy 61 KASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNCRNPNRRPWCYVQVGLK 140
Qy 121 PLVQECMVHDCADG 134
Db 141 LLVQECMVHDCADG 154
RESULT 2
Q8MHILO PRELIMINARY; PRT; 433 AA.
ID Q8MHILO
AC Q8MHILO
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN PLAU.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22155945; PubMed=12149463;
RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
RA Dicke D.A.;
RT "Increased expression of urokinase during atherosclerotic lesion
development causes arterial constriction and lumen loss, and
accelerates lesion growth."
RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY122285; AA83187.1; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65E64F36415549B0 CRC64;
Query Match 74.9%; Score 627; DB 6; Length 433;
Best Local Similarity 76.2%; Pred. No. 1.5e-59;
Matches 109; Conservative 11; Mismatches 21; Indels 2; Gaps 1;
Qy 1 SNELHGV--PSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHY 58
Db 21 SNELHGVSDASNGCCLNGGTCVTKYFSNIWRNCNCPKFGQHCCEIDTLKTCYHGDGHSY 80

```

RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RL EMBL; AK054349; BAC35743.1; -.
FT NON TER 231
SQ SEQUENCE 231 AA; 25510 MW; 25E8980A682737F2 CRC64;

Query Match 68.0%; Score 569; DB 11; Length 231;
Best Local Similarity 70.4%; Pred. No. 1.4e-53;
Matches 95; Conservative 15; Mismatches 25; Indels 0; Gaps 0;

QY 9 SNCDCLNGCTCVSNKYFSNIHWCNCPKFGGQHCHIDKSKTCYEGNGHFYRGKASTDTMG 68
DB 30 SNCGQGVGVSVYKYFSIRRCSCPKFQGEHCIDASKTCYHGNGDSYRGKANTDTKG 89

QY 69 RPLCPWNSATVLQOTVHAHRSDALQLGLGKHNYCRNPDPNRRPPWCYVQVGLKPLVQECMV 128
DB 90 RPLCAWNAVLQKPYNAHRPDAISLGLGKHNYCRNPDPNQRKRPWCYVQVGLRQFVQECMV 149

QY 129 HDCADGKFPSPPEE 143
DB 150 HDCSLSKFPSSVDQ 164

RESULT 6
Q9XT70 PRELIMINARY; PRT; 214 AA.
AC Q9XT70;
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator (fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OC NCBI_TaxId=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Yin J., Idell S.;
RT "Partial mRNA of rabbit uPA ";
RI Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC EMBL; AF097647; AAD39351.1; -.
DR HSP; P00749; LEJN.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PD000722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR GlycoProtein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
FT NON TER 214 214
SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

Query Match 40.3%; Score 337.5; DB 6; Length 214;
Best Local Similarity 78.7%; Pred. No. 1.4e-28;
Matches 59; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

QY 67 MGRCPLPWSATVLQOTVHAHRSDALQLGLGKHNYCRNPDPNRRPPWCYVQVGLKPLVQEC 126
DB 1 MDRPCLAWSANVLTITVYAHRPDALQLGLGKHNYCRNPDPHQRPPWCYVQVGLKQLIQEC 60

QY 127 MVHDCADGKKFSPSP 141

```

```

Db      61 KVHD-SSGKFPALPP 74
      ||| : ||| : ||
RESULT 7
Q9BU99 PRELIMINARY; PRT; 516 AA.
ID Q9BU99
AC Q9BU99;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; BC002795; AA02795.1; -.
DR HSP; P00750; IASH.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 2.
DR PROSITE; PS00022; Tryp_Spc; 1.
DR PROSITE; PS01186; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;

Query Match 38.5%; Score 322.5; DB 4; Length 516;
Best Local Similarity 44.4%; Pred. No. 1.6e-26;
Matches 64; Conservative 13; Mismatches 56; Indels 11; Gaps 2;

Qy 1 SNELH-----QVSNCDLGGTGVSNKYFSNIHWCNCPKFGGQHCIEDKSKTC 50
Db 23 SQEIHARPRGARSGYGGCEPRCFNGGTCCQALYFSDF-VQCPEGFAGKCEIDTRATC 81

Qy 51 YEGNGHFYEGKASTDMGRPCLPNWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDRNR 110
Db 82 YEDQGISYRGTWSTAESGAECTNWNSSALQAQPSYGRPRDAIRLGLGNHNYCRNPDR 141

Qy 111 PWCYVQVGLKPLVQECMVHDCADG 134
Db 142 PWCYVFKAGKYSEFCSTPACSEG 165

RESULT 8
Q91VP2 PRELIMINARY; PRT; 559 AA.
ID Q91VP2
AC Q91VP2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

```

```

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to plasminogen activator, tissue.
GN PLAT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; BC011256; AAH11256.1; -.
DR HSP; P00761; IANI.
DR MGD; MGI:97610; Plat.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; Inl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD000395; Kringle.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FNI; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SMC0020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 559 AA; 63122 MW; 8CCBE2BDB94514D9 CRC64;

Query Match 37.7%; Score 315.5; DB 11; Length 559;
Best Local Similarity 44.5%; Pred. No. 9.8e-26;
Matches 61; Conservative 15; Mismatches 56; Indels 5; Gaps 2;

Qy 3 ELHQP-----SNCDLNGGTGVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFY 58
Db 74 QCHSVFVRSCSEPRCFNGGTCCQALYFSDF-VQCPEGFVGRKCDIDTRATCFEEQGIT 132

Qy 59 RGKASTDMGRPCLPNWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDRNRRCVYVQVG 118
Db 133 RGTWSTAESGAECTNWNSSVLSKPYNARRPNAIKLGLGNHNYCRNPDRDLKPCWYVFA 192

Qy 119 LKPLVQECMVHDCADGK 135
Db 193 GKYTFECSTPACPKGK 209

RESULT 9
Q8SQ23 PRELIMINARY; PRT; 562 AA.
ID Q8SQ23
AC Q8SQ23;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX	NCBI_TaxID=9823;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Enamel organ;
RT	Ding Y., Xue J., Bartlett J.D.;
RA	"T-plasminogen activator in tooth tissues";
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC	-1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR	EMBL: AF346405; AM00297.1; ..
DR	HSSP: P00761; IAN.
DR	InterPro: IPR006209; EGF-like.
DR	InterPro: IPR000083; Fibrinctn1.
DR	InterPro: IPR006210; IEGF.
DR	InterPro: IPR000001; Kringle.
DR	EMBL: AF346405; AM00297.1; ..
DR	HSSP: P00761; IAN.
DR	InterPro: IPR006209; EGF-like.
DR	InterPro: IPR000083; Fibrinctn1.
DR	InterPro: IPR006210; IEGF.
DR	InterPro: IPR000001; Kringle.
DR	InterPro: IPR001254; Ser_protease_Try.
DR	Pfam: PF00008; EGF; 1.
DR	Pfam: PF00039; fnl; 1.
DR	Pfam: PF00051; kringle; 2.
DR	Pfam: PF00089; trypsin; 1.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	PRINTS; PR00018; KRINGLE.
DR	ProDom; P000395; Kringle; 2.
DR	SMART; SM00181; EGF; 1.
DR	SMART; SM00058; FN1; 1.
DR	SMART; SM00130; KR; 2.
DR	SMART; SM00020; Tryp_Spc; 1.
DR	PROSITE; PS00022; EGF_1; 1.
DR	PROSITE; PS01186; EGF_2; 1.
DR	PROSITE; PS01253; FIBRINECTIN_1; 1.
DR	PROSITE; PS00021; KRINGLE_1; 2.
DR	PROSITE; PS00070; KRINGLE_2; 2.
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.
DR	PROSITE; PS0134; TRYPSIN_HIS; 1.
DR	PROSITE; PS00135; TRYPSIN_SER; 1.
KW	EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW	Serine protease.
KW	SEQUENCE 562 AA; 63668 MW; F986B4C77CB101E8 CRC64;
SQ	
Query Match	37.2%; Score 311.5; DB 6; Length 562;
Best Local Similarity	43.5%; Pred. No. 2.7e-25;
Matches	60; Conservative 15; Mismatches 58; Indels 5; Gaps 3;
QY	3 ELHQP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFY 58
DB	77 QCHSVPVKSCSEPRCNGTCLQAIYFSDP-VCCQPEVFGIQCIEDARATCYEDQGITY 135
QY	59 RGKASTDTMGRCPLPWNATVLOQTYHAHRSALQLGLGKHNYCRPNRRPWCYQVG 118
DB	136 RGTWSTTESGAECVNNWNTSGLASMPYNGRRPDAVKLGLNHNHYCRNPDKDSKPCWYIFKA 195
QY	119 LKPLVQECMVHDCAGKK 136
DB	196 EKYSDFECSTPACTKEKE 213
RESULT 10	
Q8MKB1	PRELIMINARY; PRT; 564 AA.
ID	Q8MKB1
AC	Q8MKB1;
DT	01-OCT-2002 (TrEMBLrel. 22, Created)
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Tissue-type plasminogen activator.
OS	Oryctolagus cuniculus (Rabbit).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX	NCBI_TaxID=9986;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC	-1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR	EMBL: AY029518; AAK40240.1; ..
DR	InterPro: IPR000001; Kringle.
DR	Pfam; PF00008; EGF; 2.
DR	Pfam; PF00051; kringle; 1.
DR	Pfam; PF00089; trypsin; 1.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	InterPro: IPR001254; Ser_protease_Try.
DR	Pfam; PF00008; EGF; 2.
DR	Pfam; PF00051; kringle; 1.
DR	Pfam; PF00089; trypsin; 1.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	InterPro: IPR001254; Ser_protease_Try.
DR	Pfam; PF00008; EGF; 1.
DR	Pfam; PF00039; fnl; 1.
DR	Pfam; PF00051; kringle; 2.
DR	Pfam; PF00089; trypsin; 1.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	PRINTS; PR00018; KRINGLE.
DR	ProDom; P000395; Kringle; 2.
DR	SMART; SM00181; EGF; 1.
DR	SMART; SM00058; FN1; 1.
DR	SMART; SM00130; KR; 2.
DR	SMART; SM00020; Tryp_Spc; 1.
DR	PROSITE; PS00022; EGF_1; 1.
DR	PROSITE; PS01186; EGF_2; 1.
DR	PROSITE; PS01253; FIBRINECTIN_1; 1.
DR	PROSITE; PS00021; KRINGLE_1; 2.
DR	PROSITE; PS00070; KRINGLE_2; 2.
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.
DR	PROSITE; PS0134; TRYPSIN_HIS; 1.
DR	PROSITE; PS00135; TRYPSIN_SER; 1.
KW	EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW	Serine protease.
KW	SEQUENCE 564 AA; 62726 MW; 459D8BAC6D4A937C CRC64;
SQ	
Query Match	36.0%; Score 301.5; DB 6; Length 564;
Best Local Similarity	48.7%; Pred. No. 3.2e-24;
Matches	56; Conservative 13; Mismatches 41; Indels 5; Gaps 3;
QY	5 HOVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB	80 HSPVQSCSEPRCLNGGTCQALYFSDP-VCCQPEGVGRKEVDTRARCYEDRGIVRG 138
QY	61 KASTDTMGRCPLPWNATVLOQTYHAHRSALQLGLGKHNYCRPNRRPWCYV 115
DB	139 TWSSTTESGAQCYNWNNWTLALKPKYSGRKNALRLGLGNHNHYCRNPDRTKPCYV 193
RESULT 11	
Q8KOD2	PRELIMINARY; PRT; 517 AA.
ID	Q8KOD2
AC	Q8KOD2;
DT	01-OCT-2002 (TrEMBLrel. 22, Created)
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Hypothetical protein.</

```

DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; EGF like.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp SPG; 1.
DR PROSITE; PS00022; EGF 1; 3.
DR PROSITE; PS01186; EGF 2; 2.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00021; KRINGLE 2; 1.
DR PROSITE; PS00070; TRYPSIN_DOM; 1.
DR PROSITE; PS00440; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 517 AA; 57326 MW; 3855A42035A5EA59 CRC64;

Query Match 31.7%; Score 265; DB 11; Length 517;
Best Local Similarity 41.1%; Pred. No. 2.6e-20;
Matches 53; Conservative 19; Mismatches 51; Indels 6; Gaps 4;

QY 13 CLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
DB 116 CQNGGVCSHRRSRF-TCACPDQYKGFCEIGPD-DCVVGDGYSYRGKSKTVNQNECL 173

QY 73 PWSNATVLOQTYHAHRSALQLGLGKHNYCRNPNRRPWCYVQVGLKPLVQE-CMYHDC 131
DB 174 YWNSHLLLOETYNMFEMEDAETHGIAHNFRCNPDGDHPCWCFKYNSEKVKWEYCDVTVC 233

QY 132 ADGKXPSP 140
DB 234 ---PVPDTP 239

RESULT 12
Q14520
ID Q14520 PRELIMINARY; PRT; 560 AA.
AC Q14520; O00663.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE HGF activator like protein (Hyaluronan binding protein 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kitamura N.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96425001; PubMed=8827452;
RA Choi-Miura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T.,
RA Tomita M.;
RT "Purification and characterization of a novel hyaluronan-binding
RT protein (PHBP) from human plasma: it has three EGF, a kringle and a
RT serine protease domain, similar to hepatocyte growth factor
RT activator."
RL J. Biochem. 119:1157-1165(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon, and Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR ENBL; D49742; BAA08576.1;
DR ENBL; S83182; AAB46909.1;
DR ENBL; BC031412; AAB31412.1;
DR HSSP; P00763; IDPO.
DR MEROPS; S01.033; -.
DR Genew; HGNC:4798; HABP2.
DR InterPro; IPR001314; Chymotrypsin.

```

```

DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR PROSITE; PS00022; EGF 1; 3.
DR PROSITE; PS01186; EGF 2; 2.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00021; KRINGLE 2; 1.
DR PROSITE; PS00440; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 560 AA; 62671 MW; 5C1907230784ACD4 CRC64;

Query Match 31.5%; Score 263.5; DB 4; Length 560;
Best Local Similarity 41.0%; Pred. No. 4.1e-20;
Matches 55; Conservative 18; Mismatches 56; Indels 5; Gaps 4;

QY 13 CLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
DB 159 CQNGATCSRHKRSKF-TCACPDQYKGFCEIG-SDCYVGDGYSYRGKRNRTVQNHACL 216

QY 73 PWSNATVLOQTYHAHRSALQLGLGKHNYCRNPNRRPWCYVQVGLKPLVQE-CMYHDC 131
DB 217 YWNSHLLLOETYNMFEMEDAETHGIGEHNFRCNPDADKPCFKIVTKNDKWEYCDVSAC 276

QY 132 A--DGKXPSPPEE 143
DB 277 SAQDVAYPEESPT 290

RESULT 13
Q97507
ID Q97507 PRELIMINARY; PRT; 616 AA.
AC Q97507;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE FXII.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Takahashi T., Kihara T.;
RT "Porcine liver factor XII."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AB022426; BAA37148.1;
DR HSSP; P00763; IDPO.
DR MEROPS; S01.211; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR006210; IBGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.

```

```

DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0013; FNTYPEII.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000995; FN Type II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS01253; FIBRONECTIN_2; 1.
DR PROSITE; PS00023; KRINGLE_1; 1.
DR PROSITE; PS00021; KRINGLE_2; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00070; TRYPSIN_DOM; 1.
DR PROSITE; PS02040; TRYPSIN_HIS; 1.
DR PROSITE; PS00134; TRYPSIN_SER; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase;
KW Serine protease.
SQ SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBED1A9 CRC64;

Query Match 30.8%; Score 258; DB 6; Length 616;
Best Local Similarity 39.0%; Pred.No.1.8e-19;
Matches 57; Conservative 15; Mismatches 58; Indels 16; Gaps 5;

Qy 6 QVPSNCDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRGKASTD 65
Db 176 QVCSNTNCLNGSGCLQTE---GHLRCRCPTGYAGRLCDVLRKCYSDRGLSVGMAQTT 232
Qy 66 TMGRPCLPWNSATVLOQTY-HAHSRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLKPLVQ 124
Db 233 LSGAPCPQWAS---EATYWNNTAEQALNWLGLGHAFRCNPDNDTRPWCFFWRGDQLSWQ 288
Qy 125 ECMVHDCAD--GKKP-----SSPPE 142
Db 289 YCLARCOAPTEGAFFILTPQSPSE 314

RESULT 14
Q8VCS4 PRELIMINARY; PRT; 653 AA.
ID Q8VCS4
AC Q8VCS4;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical 70.6 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP TISSUE=Liver;
RC STRAUSBERG R.;
RA Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAH19376.1; -.
DR HSP; P00761; IAN1.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000883; Fibrin.
DR InterPro; IPR000562; FN Type_II.
DR InterPro; IPR000562; FN Type_II.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.

```

```

DR PRINTS; PRO0089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0013; FNTYPEII.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000995; FN Type II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match 29.9%; Score 250.5; DB 11; Length 653;
Best Local Similarity 42.0%; Pred.No.1.2e-18;
Matches 50; Conservative 11; Mismatches 41; Indels 17; Gaps 2;

Qy 5 HQVPSNCDCLNGTGVSNKYFNSNIHW-----CNCPKFGGQHCEIDKSKTCYEGNGH 56
Db 239 HTACLSSPCLNGTGC-----HLIVGTGTSVCTCPGLYAGRFNCNIVPTHECFPLNGT 289
Qy 57 FYRGKASTDTMGRPCLPWNSATVLOQTYHAHSRSDALQLGLGKHNYCRPNDRRRPWCYV 115
Db 290 EYRGVASTAASGLSCLAMNSDLLYQELHVDVSAVAALLGLGPHAYCRNPKDPRWCYV 348

RESULT 15
Q35727 PRELIMINARY; PRT; 597 AA.
ID Q35727
AC Q35727;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Factor XII.
GN F12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP TISSUE=Liver;
RC SCHLOESSER M., Schwager S., Engel W.;
RA Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; X95571; CAA67891.1; -.
DR HSP; P00760; IAO7.
DR HSP; P00761; IAN1.
DR MGD; MGI1891012; F12.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000883; Fibrin.
DR InterPro; IPR000562; FN Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.

```

```

DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR DR PROSITE; PS00021; KRINGLE_1; 1.
DR DR PROSITE; PS00070; KRINGLE_2; 1.
DR DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KMGlycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON_TER 395
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 27.8%; Score 233; DB 4; Length 395;
Best Local Similarity 41.2%; Pred. No. 5,7e-17;
Matches 47; Conservative 12; Mismatches 49; Indels 6; Gaps 3;

Qy 30 WCNCPKFGQGHCEIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNASATVLOOTYHAHRS 89
Db 33 WCNS-----GRAQCS-EGNSDCYFENGSAVSGTHTSLTESGASCLPWNISMILIGKVYTAQNP 87

Qy 90 DALQLGLGKHNYCRNPNRRPPCYQVGLKPLVQECMVHDCAD-GKKPSSPPE 142
Db 88 SAQALGLGKHNYCRNPDGDGAKPKHVLKWRLLTWYCYDVPSCSTCGLRQYSQFQ 141

RESULT 17
Q8I2Z5 PRELIMINARY; PRT; 615 AA.
AC Q8I2Z5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Coagulation factor XII-Mie.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wada H., Nishioaka J., Nakatani K., Kasei Y., Abe Y., Nobori T.;
RT "Molecular characterization of coagulation factor XII-Mie.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB095845; BAC23095.1; -.
DR SWBL;
SQ SEQUENCE 615 AA; 67735 MW; 030508870A0C7EDB CRC64;

Query Match 27.8%; Score 233; DB 4; Length 615;
Best Local Similarity 42.3%; Pred. No. 9,2e-17;
Matches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;

Qy 13 CLNGGTCVSNKYPSNTHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
Db 183 CLHGGRCLE---VEGHRLLCHCPGVGTGPFCDVDTKASCDVGRGLSYRGLARTLTLSGAPCQ 239

Qy 73 PWNASATVLOOTY-HAHRSDALQLGLGKHNYCRNPNRRPPCYV 115
Db 240 PNAS-----EATYRNVTAEQARNWGLGGAFCRNPDNDIRPWCVF 279

RESULT 18
Q8NG20 PRELIMINARY; PRT; 90 AA.
AC Q8NG20;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen/activator kringle.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "Production of kringle fragment.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

```



```

CC  -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR  EMBL; AF282882; AAM52248.1; -.
DR  InterPro; IPR000001; Kringle.
DR  Pfam; PF00051; kringle; 1.
DR  PRINTS; PR00018; KRINGLE.
DR  ProDom; PD000395; Kringle; 1.
DR  SMART; SM00130; KR_1.
DR  PROSITE; PS00021; KRINGLE_1; 1.
DR  PROSITE; PS00070; KRINGLE_2; 1.
KW  Glycoprotein; Kringle.
SQ  SEQUENCE 90 AA; 9804 MW; A33887F9FDF4C7B1 CRC64;

Query Match 26.1%; Score 218.5; DB 4; Length 90;
Best Local Similarity 49.4%; Pred. No. 4.2e-16;
Matches 41; Conservative 5; Mismatches 36; Indels 1; Gaps 1;

QY 50 CYEGNGHFYRGKASTDTMGSPCLPWN SATVLQOQYHAHRSDALQLGKKNYCRNPNR 109
Db 8 CYFGNGSAVRGTHSLTSGASCLPWN SMILIGKVYTAQNSAQAALGLGKKNYCRNP 67
QY 110 RPKCYVQVGLKPLVQECMWHDC A 132
Db 68 KPWCYT-TNPRKLYDYCDVPOCA 89

RESULT 19
Q90675 PRELIMINARY; PRT; 202 AA.
ID AC Q90675
RC STRAIN=White Leghorn;
RX MEDLINE=97199025; PubMed=9047000;
RA Johnson A.L., Bridgham J.T., Anthony R.V.;
RT "Expression of avian urokinase and tissue-type plasminogen activator
RL messenger ribonucleic acid during follicle development and atresia.";
RL Biol. Reprod. 56:581-588(1997).
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR  EMBL; U31986; AAA74955.1; -.
DR  HSPSP; P00750; IRTF.
DR  MEROPS; S01.232; -.
DR  InterPro; IPR000001; Kringle.
DR  InterPro; IPR001254; Ser. protease_Try.
DR  Pfam; PF00051; kringle; 2.
DR  Pfam; PF00083; trypsin; 1.
DR  PRINTS; PR00018; KRINGLE.
DR  ProDom; PD000395; Kringle; 2.
DR  SMART; SM00130; KR; 1.
DR  PROSITE; PS00021; KRINGLE_1; 1.
DR  PROSITE; PS00070; KRINGLE_2; 2.
DR  PROSITE; PS50240; TRYPSIN_DOM; 1.
DR  PROSITE; PS00134; TRYPSIN_HIS; 1.
KW  Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT  NON TER 1
FT  NON TER 202
SQ  SEQUENCE 202 AA; 23148 MW; 049DB42941D83AFC CRC64;

Query Match 23.6%; Score 197.5; DB 13; Length 202;
Best Local Similarity 36.9%; Pred. No. 1.9e-13;
Matches 41; Conservative 11; Mismatches 50; Indels 9; Gaps 2;

QY 30 WNCNCFK--KFGQSQCE-----IDKSKTCYEGNGHFYRGKASTDTMGSPCLPWN SATVL 80

```

RESULT	23
Q8WNR1	PRELIMINARY; PRT; 359 AA.
ID	Q8WNR1
AC	01-MAR-2002 (TRENBLrel. 20, Created)
DT	01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT	01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE	Plasmimogen (Fragment).
DE	Canis familiaris (Dog).
OS	Canis familiaris (Dog).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX	NCBI TaxID=9615.

RP SEQUENCE FROM N.A.
RA Pirie-Shepherd S.R., Coffman K.T., Resnick D., Chan R., Kisker C.,
RT Folkman J., Waters D.J.;
RT "Angiostatin is Detectable in the Urine of Dogs with Spontaneous Bone
RT Cancer";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
EMBL; AY069985; AAL58519.1; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003966; Prothrombin.
DR Pfam; PF00051; kringle; 4.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 4.
DR

AC	Q9ROW3;
AD	01-MAY-2000 (TReMBLrel. 13, Created)
AE	01-MAY-2000 (TReMBLrel. 13, Last sequence update)
AF	01-MAY-2000 (TReMBLrel. 23, Last annotation update)
AG	01-MAR-2003 (TReMBLrel. 23, Last annotation update)
AH	Plasminogen protein precursor (EC 3.4.21.7).
AI	PLASMINOGEN.
AJ	GN
AK	Rattus norvegicus (Rat).
AL	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AM	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
AN	NCBI_TaxID=10116;
AO	[1]
AP	SEQUENCE FROM N.A.
AQ	TISSUE=Liver;
AR	Bangert K., Johnsen A.H., Thorsen S.;
AS	"Rat plasminogen: cDNA and gene structure."
AT	Submitted (MAY-1999) to the EMBL/GenBank/DBP1 databases

```

RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Liver;
RX  MEDLINE=91250378; PubMed=1645711;
RA  Kanalas J.J.; Makker S.P.;
RT  "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RL  receptor site for plasminogen."
RJ  J. Biol. Chem. 266:10825-10829(1991).
CC  -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR  EMBL; AJ242649; CAB46014.1; -.
DR  HSSP; P00747; 1PMK.
DR  InterPro; IPR001314; Chymotrypsin.
DR  InterPro; IPR000001; Kringle.
DR  InterPro; IPR003014; PAN.
DR  InterPro; IPR003609; Pan app.
DR  InterPro; IPR003966; Prothrombin.
DR  InterPro; IPR001254; Ser. protease Try.
DR  InterPro; IPR001400; Somatotropin.
DR  Pfam; PF00051; kringle; 5.
DR  Pfam; PF00024; PAN; 1.
DR  PRINTS; PR00722; CHYMOTRYPSIN.
DR  PRINTS; PR00018; KRINGLE.
DR  PRINTS; PR01505; PROTHROMBIN.
DR  ProDom; PD000395; Kringle; 5.
DR  SMART; SM00130; KR; 4.
DR  SMART; SM00473; PAN_AP; 1.
DR  SMART; SM00020; TRYPSIN; 1.
DR  PROSITE; PS00021; KRINGLE_1; 5.
DR  PROSITE; PS00070; KRINGLE_2; 5.
DR  PROSITE; PS00338; SOMATOTROPIN_2; 1.
DR  PROSITE; PS02400; TRYPSIN_DOM; 1.
DR  PROSITE; PS00134; TRYPSIN_HIS; 1.
DR  PROSITE; PS00135; TRYPSIN_SER; 1.
KW  Glycoprotein; Hydrolase; Kringle; Protease; Serine protease; Signal.
FT  SIGNAL 1 19
FT  CHAIN 20 812 PLASMINOGEN.
SQ  SEQUENCE 812 AA; 90535 MW; 8C703C51410EBC9E CRC64;

Query Match 19.8%; Score 165.5; DB 11; Length 812;
Best Local Similarity 27.8%; Pred. No. 2.5e-09;
Matches 50; Conservative 22; Mismatches 51; Indels 57; Gaps 12;

QY 5 H0VPSNCDCLNGTGVSNKYFSN-----IHW-C-NCPKFGGHCIE-----DKS-- 47
DB 308 NRTPEFPCN-----LEENYCRNPDCETAPWCYTDSQLRWYCIIPSCGSSVSPDQSDS 363
QY 48 -----KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOITYHAHRSDALQL 94
DB 364 SVLPETQPVVQECYQNGKSYRGSTSTNTGKKCSW-----VSMTPHSHSKTANFPDA 418
QY 95 GLGKYNPCNPN-RRPWCYVQVGLKPLV--QECNVHDCAD-----GKKPSP 140
DB 419 GL-EMNYCRNPNDQRPWCFT---TDPVSRWEYCNLKKCSETGGGVAESAIVPQVSPAP 474

RESULT 25
ID Q90Y90 PRELIMINARY; PRT; 452 AA.
AC Q90Y90;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Similar to receptor tyrosine kinase-like orphan receptor 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC006374; AA06374.1; -.
DR HSSP; P00747; ICEA.
DR InterPro; IPR000024; Fz_domain.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00038; FZ; 1.
DR PROSITE; PS00835; IG LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
RX MEDLINE=21167372; PubMed=11267660;

```

```

RA Nakamura T., Aoki S., Kitajima K., Funakoshi H., Takahashi T.,
RP Matsumoto K., Nakamura T.;
RT "Molecular cloning and characterization of Kremen, a novel kringle-
RL containing transmembrane protein."
RJ Biochim. Biophys. Acta 1518:63-72(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 CUB DOMAIN.
DR EMBL; AB070851; BAB64294.1; -.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR003001; Kringle.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
SQ SEQUENCE 452 AA; 50188 MW; ED24BCDIAF4564E2 CRC64;

Query Match 19.0%; Score 159; DB 13; Length 452;
Best Local Similarity 40.8%; Pred. No. 6.7e-09;
Matches 31; Conservative 9; Mismatches 32; Indels 4; Gaps 2;

QY 42 CEIDKSKTCYEGNGHFYRGKASTDTM--GRPCLPWSATVLQOITYHAHRSDALQLGLGKH 99
DB 22 CSDSFHSECTVNGADVRGTQNTSLDGGKPCLPWNE--TFQHPYNTLKYPNGGGLGEH 79
QY 100 NYCRRNPDRRRPWCYV 115
DB 80 NYCRRNPQGDVPCYI 95

RESULT 26
ID Q9BRB6 PRELIMINARY; PRT; 393 AA.
AC Q9BRB6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to receptor tyrosine kinase-like orphan receptor 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC006374; AA06374.1; -.
DR HSSP; P00747; ICEA.
DR InterPro; IPR000024; Fz_domain.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00038; FZ; 1.
DR PROSITE; PS00835; IG LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.

```



```
SQ SEQUENCE 454 AA; 50041 MW; 974E30744C187B2F CRC64;

Query Match      18.8%; Score 158; DB 6; Length 454;
Best Local Similarity 39.8%; Pred. No. 8.6e-09;
Matches 35; Conservative 10; Mismatches 31; Indels 12; Gaps 4;

QY 50 CYEGNGHFYRGKASTDTMGRPCLPWNSATVQ--QTYHAHRSALQLGLGKHNYCRNP 107
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 102 CYHGDQOSYRGSTTVTGRTCQSWSMTPHQHKRTPENHPNDGLTM-----NYCENPDA 156

QY 108 RRPWCYVQVGLKPLV--QECMVHDCAD 133
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 157 DTGPGWCFT---MDPSVRWEYCNLTRGSD 181

RESULT 30
Q9UIRS PRELIMINARY; PRT; 113 AA.
AC Q9UIRS;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21181705; PubMed=11285247;
RA Ogorelikova M., Kraft H.G., Enholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringle IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158863; AAF03680.1; -.
DR EMBL; AF158862; AAF03680.1; JOINED.
DR HSSP; P00747; LPMK.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SMO0130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PSS0070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12685 MW; F3D65681D9B5253A CRC64;

Query Match      18.8%; Score 157; DB 4; Length 113;
Best Local Similarity 34.8%; Pred. No. 2.4e-09;
Matches 39; Conservative 13; Mismatches 36; Indels 24; Gaps 6;

QY 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATV--LQTYHAHRSALQLGLGKHNYCRNP 105
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 9 RCYHNGQSYRGSTTVTGRTCQSWSMTPHQHKRTPENHPNDGLTM-----NYCRNP 63

QY 106 DNRPRPCYVQVGLKPLV--QECMVHDCADG-----KKPS-SPPEE 143
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 64 DADTGPWCFT---MDPSIRWEYCNLTRGSDTEGTVAAPTIVQPSLGPDPSE 112
```

Search completed: December 3, 2003, 14:43:00
Job time : 18.5948 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:03:04 ; Search time 14.681 Seconds
(without alignments)
1811.566 Million cell updates/sec

Title: US-09-880-503-8
Perfect score: 837
Sequence: 1 SNELHQPVSNCDCGCTCV.....QECMVHDCADGKXSSPPEE 143

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata1/pubpaa/FCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata1/pubpaa/CTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/prodata1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/prodata1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	837	100.0	143	US-09-880-503-8	Sequence 8, Appli
2	837	100.0	337	US-10-108-698-6266	Sequence 6266, Ap
3	837	100.0	411	US-09-880-503-3	Sequence 3, Appli
4	837	100.0	431	US-09-264-468B-1	Sequence 1, Appli
5	837	100.0	431	US-10-301-822-161	Sequence 161, App
6	837	100.0	431	US-10-131-985-21	Sequence 21, Appli
7	837	100.0	431	US-10-076-421-2	Sequence 2, Appli
8	837	100.0	431	US-10-171-311-184	Sequence 184, App
9	837	100.0	431	US-10-193-656-4	Sequence 4, Appli
10	835	99.8	431	US-10-247-671-149	Sequence 149, App
11	827	98.8	411	US-10-407-821-2	Sequence 2, Appli
12	793	94.7	135	US-09-880-503-4	Sequence 4, Appli
13	793	94.7	138	US-09-984-186-12	Sequence 12, Appli
14	793	94.7	138	US-10-237-687-12	Sequence 12, Appli
15	793	94.7	138	US-10-237-708-12	Sequence 12, Appli

16	793	94.7	138	15	US-10-237-866-12	Sequence 12, Appli
17	793	94.7	138	15	US-10-237-871-12	Sequence 12, Appli
18	793	94.7	138	15	US-10-237-624-12	Sequence 12, Appli
19	793	94.7	403	10	US-09-880-503-6	Sequence 6, Appli
20	554	66.2	96	10	US-09-880-503-9	Sequence 9, Appli
21	510	60.9	88	10	US-09-880-503-1	Sequence 1, Appli
22	328.5	39.2	527	11	US-09-987-457-18	Sequence 18, Appli
23	328.5	39.2	527	11	US-09-987-455-19	Sequence 19, Appli
24	328.5	39.2	527	9	US-09-959-271-7	Sequence 7, Appli
25	328.5	39.2	562	10	US-09-974-298-145	Sequence 145, App
26	328.5	39.2	562	12	US-10-443-701-4	Sequence 4, Appli
27	328.5	39.2	562	15	US-10-193-656-8	Sequence 8, Appli
28	293	35.0	49	12	US-10-349-543-5	Sequence 5, Appli
29	288	34.4	323	10	US-09-880-503-7	Sequence 7, Appli
30	283	31.8	47	10	US-09-880-503-2	Sequence 2, Appli
31	268	32.0	44	12	US-10-349-543-1	Sequence 1, Appli
32	263.5	31.5	560	10	US-09-912-559-3	Sequence 3, Appli
33	263.5	31.5	560	10	US-09-912-559-4	Sequence 4, Appli
34	263.5	31.5	560	15	US-10-172-712-32	Sequence 32, Appli
35	262	31.3	655	15	US-10-172-712-28	Sequence 28, Appli
36	233	27.8	615	11	US-09-858-909-2	Sequence 2, Appli
37	233	27.8	615	15	US-10-172-712-30	Sequence 30, Appli
38	224	26.8	326	14	US-10-057-951-3	Sequence 3, Appli
39	224	26.8	372	9	US-09-084-491A-3	Sequence 3, Appli
40	224	26.8	372	14	US-10-102-704-3	Sequence 3, Appli
41	218.5	26.1	354	11	US-09-987-457-10	Sequence 10, Appli
42	218.5	26.1	354	11	US-09-987-455-11	Sequence 11, Appli
43	218.5	26.1	377	11	US-09-987-455-8	Sequence 8, Appli
44	179.5	21.4	343	11	US-09-987-457-14	Sequence 14, Appli
45	179.5	21.4	343	11	US-09-987-455-15	Sequence 15, Appli

ALIGNMENTS

RESULT 1

- US-09-880-503-8
- ; Sequence 8, Application US/09880503
- ; Patent No. US20020131964A1
- ; GENERAL INFORMATION:
- ; APPLICANT: CINES, Douglas B
- ; APPLICANT: HIGAZI, Abd Al-Root
- ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND TISSUE CONTRACTILITY
- ; FILE REFERENCE: 9596-331
- ; CURRENT APPLICATION NUMBER: US/09/880,503
- ; CURRENT FILING DATE: 2001-06-13
- ; PRIOR APPLICATION NUMBER: US 60/212,847
- ; PRIOR FILING DATE: 2000-06-20
- ; NUMBER OF SEQ ID NOS: 18
- ; SOFTWARE: PatentIn Ver. 2.1
- ; SEQ ID NO 8
- ; TYPE: PRT
- ; LENGTH: 143
- ; ORGANISM: Homo sapiens
- US-09-880-503-8

Query Match	100.0%;	Score	837;	DB	10;	Length	143;
Best Local Similarity	100.0%;	Pred. No.	1.7e-76;	Mismatches	0;	Gaps	0;
Matches	143;	Conservative	0;	Indels	0;		
Qy	1	SNELHQPVSNCDCGCTCVSNKYFNSIHWNCPCFKFGQHCIEDKSKTCYEGNGHFYRG	60				
Db	1	SNELHQPVSNCDCGCTCVSNKYFNSIHWNCPCFKFGQHCIEDKSKTCYEGNGHFYRG	60				
Qy	61	KASDTTMRPCLPNSATVLOQTYHAHRSALQJGLGKHNYCRNPDNRRRPWCYVQVLK	120				
Db	61	KASDTTMRPCLPNSATVLOQTYHAHRSALQJGLGKHNYCRNPDNRRRPWCYVQVLK	120				
Qy	121	PLVQECMVHDCADGKXSSPPEE	143				
Db	121	PLVQECMVHDCADGKXSSPPEE	143				

RESULT 2
US-10-106-698-6266
; Sequence 6266, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005F1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patent In Ver. 3.0
; SEQ ID NO 6266
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-6266
Query Match 100.0%; Score 837; DB 15; Length 337;
Best Local Similarity 100.0%; Pred. No. 4.2e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHFYRG 60
DB 27 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHFYRG 86
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
DB 87 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNCRNPNRRRPPWCYVQVGLK 146
QY 121 PLVQECMVHDCADGKPSPPPEE 143
DB 147 PLVQECMVHDCADGKPSPPPEE 169
RESULT 3
US-09-880-503-3
; Sequence 3, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Rouf
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-3
Query Match 100.0%; Score 837; DB 10; Length 411;
Best Local Similarity 100.0%; Pred. No. 5.2e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120

DB 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPSPPPEE 143
DB 121 PLVQECMVHDCADGKPSPPPEE 143
RESULT 4
US-09-264-468B-1
; Sequence 1, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310.US.F1
; CURRENT APPLICATION NUMBER: US/09/264,468B
; CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SIGNAL
; LOCATION: (1)...(20)
; OTHER INFORMATION: Leader sequence
; NAME/KEY: VARIANT
; LOCATION: (279)...(279)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: VARIANT
; LOCATION: (302)...(302)
; OTHER INFORMATION: Xaa = any amino acid
US-09-264-468B-1
Query Match 100.0%; Score 837; DB 10; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.5e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCCEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNCRNPNRRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKPSPPPEE 143
DB 141 PLVQECMVHDCADGKPSPPPEE 163
RESULT 5
US-10-301-822-161
; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert

APPLICANT: Monahan, John E.
APPLICANT: Thibodeau, Stephen N.
APPLICANT: Burgart, Lawrence J.
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF COLON CANCER
FILE REFERENCE: MP01-029P2NM
CURRENT APPLICATION NUMBER: US/10/301,822
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 60/339,971
PRIOR FILING DATE: 2001-12-10
PRIOR APPLICATION NUMBER: US 60/361,978
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/381,988
PRIOR FILING DATE: 2002-05-20
NUMBER OF SEQ ID NOS: 228
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 161
LENGTH: 431
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-301-822-161

Query Match 100.0%; Score 837; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.5e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
Qy 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 140

Qy 121 PLVQECMVHDCADGKPKSPPEE 143
Db 141 PLVQECMVHDCADGKPKSPPEE 163

RESULT 6
US-10-131-985-21
Sequence 21, Application US/10131985
Publication No. US2003019940A1
GENERAL INFORMATION:
APPLICANT: Davies, Michael J
APPLICANT: Fish, Paul V
APPLICANT: Huggins, Jonathan P
APPLICANT: McIntosh, Fraser S
APPLICANT: Occleston, Nicholas L
TITLE OF INVENTION: Composition
FILE REFERENCE: PCS 10391A
CURRENT APPLICATION NUMBER: US/10/131,985
PRIOR FILING DATE: 2002-04-25
PRIOR APPLICATION NUMBER: US/09/726,295
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: GB 9930768.8
PRIOR FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatenIn Ver. 2.1
SEQ ID NO 21
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-985-21

Query Match 100.0%; Score 837; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.5e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60

Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
Qy 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 140
Qy 121 PLVQECMVHDCADGKPKSPPEE 143
Db 141 PLVQECMVHDCADGKPKSPPEE 163

RESULT 7
US-10-076-421-2
Sequence 2, Application US/10076421
Publication No. US20020193304A1
GENERAL INFORMATION:
APPLICANT: WADA, MANABU
APPLICANT: WADA, NAOKO
TITLE OF INVENTION: ANTI-HIV AGENTS
FILE REFERENCE: HAYAK-9
CURRENT APPLICATION NUMBER: US/10/076,421
CURRENT FILING DATE: 2002-05-17
PRIOR APPLICATION NUMBER: JP 2001-42655
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: JP 2001-184284
PRIOR FILING DATE: 2001-06-19
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
US-10-076-421-2

Query Match 100.0%; Score 837; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.5e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
Qy 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 140
Qy 121 PLVQECMVHDCADGKPKSPPEE 143
Db 141 PLVQECMVHDCADGKPKSPPEE 163

RESULT 8
US-10-171-311-184
Sequence 184, Application US/10171311
Publication No. US20030087270A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Chen, Yan
APPLICANT: Zhao, Xumei
APPLICANT: Monahan, John
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Giatt, Karen
APPLICANT: Gannavarapu, Manjula
APPLICANT: Hoerish, Sebastian
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
TITLE OF INVENTION: OF CERVICAL CANCER
FILE REFERENCE: MRI-035
CURRENT APPLICATION NUMBER: US/10/171,311
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: US 60/298,159
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,155

; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 184
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-184

Query Match 100.0%; Score 837; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.5e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
DB 81 KASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGKPSPPPEE 143
DB 141 PLVQECMVHDCADGKPSPPPEE 163

RESULT 9
US-10-193-656-4
; Sequence 4, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: Li, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3910/1J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,192
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00749
; DATABASE ENTRY DATE: 1996-07-21
; RELEVANT RESIDUES: (1)..(431)
US-10-193-656-4

Query Match 100.0%; Score 837; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.5e-76;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
DB 81 KASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGKPSPPPEE 143
DB 141 PLVQECMVHDCADGKPSPPPEE 163

RESULT 10

US-10-247-671-149
; Sequence 149, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1
US-10-247-671-149

Query Match 99.8%; Score 835; DB 12; Length 431;
Best Local Similarity 99.3%; Pred. No. 8.7e-76;
Matches 142; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
DB 81 KASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGKPSPPPEE 143
DB 141 PLVQECMVHDCADGKPSPPPEE 163

RESULT 11
US-10-407-821-2
; Sequence 2, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS
; FILE REFERENCE: UTSN:022US
; CURRENT APPLICATION NUMBER: US/10/407,821
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-821-2

Query Match 98.8%; Score 827; DB 12; Length 411;
Best Local Similarity 99.3%; Pred. No. 5.3e-75;
Matches 142; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNELHVPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHVPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPPSPPEE 143
DB 121 LLVQECMVHDCADGKPPSPPEE 143

RESULT 12
US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-4

Query Match 94.7%; Score 793; DB 10; Length 135;
Best Local Similarity 100.0%; Pred. No. 4.1e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHVPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHVPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

RESULT 13
US-09-884-186-12
; Sequence 12, Application US/09884186
; Patent No. US2002015101A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/884,186
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-884-186-12

Query Match 94.7%; Score 793; DB 10; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHVPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 4 SNELHVPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 63
C 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
DB 64 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 14
US-10-237-667-12
; Sequence 12, Application US/10237667
; Publication No. US20030022308A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,667
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-667-12

Query Match 94.7%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDLNGGTCVSNKYFNSNHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 4 SNELHQVPSNCDLNGGTCVSNKYFNSNHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRCPLPWSNATVLQOTYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 120
DB 64 KASTDTMGRCPLPWSNATVLQOTYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 15
US-10-237-708-12
Sequence 12, Application US/10237708
Publication No. US20030036170A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,708
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-708-12

Query Match 94.7%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDLNGGTCVSNKYFNSNHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 4 SNELHQVPSNCDLNGGTCVSNKYFNSNHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRCPLPWSNATVLQOTYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 120
DB 64 KASTDTMGRCPLPWSNATVLQOTYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 16
US-10-237-866-12
Sequence 12, Application US/10237866
Publication No. US20030036171A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh

OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,866
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D. Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12

Query Match 94.7%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 4 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPDNRRRRCWYVQVGLK 120
Db 64 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPDNRRRRCWYVQVGLK 123
QY 121 PLVQECMVHDCADGX 135
Db 124 PLVQECMVHDCADGX 138

RESULT 17
US-10-237-871-12
Sequence 12, Application US/10237871
Publication No. US20030036172A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1

SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,871
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D. Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-871-12
Query Match 94.7%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 4 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPDNRRRRCWYVQVGLK 120
Db 64 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPDNRRRRCWYVQVGLK 123
QY 121 PLVQECMVHDCADGX 135
Db 124 PLVQECMVHDCADGX 138
RESULT 18
US-10-237-624-12
Sequence 12, Application US/10237624
Publication No. US20030082747A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)


```

RESULT 25
US-09-974-298-145
; Sequence 145, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10

```

; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 145
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CD1
US-09-974-298-145

Query Match 39.2%; Score 328.5; DB 10; Length 562;
Best Local Similarity 46.3%; Pred. No. 9.3e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;
QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHGY 58
DB 77 QCHSVPVKSCSEPRCFNGGTCQALYFSDP-VQCQPEGFAGKCCIEDTRATCYEDQGISY 135
QY 59 RKGASTDTMGRCPLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVG 118
DB 136 RGTWSTAESGAECTNWNSSALAKPYSGRRPDARLGLGNHNYCRNPDNRDRSKPWCYVFK 195
QY 119 LKPLVQECMVHDCADG 134
DB 196 GKYSEFCSTPACSEG 211

RESULT 26
US-10-443-701-4
; Sequence 4, Application US/10443701
; Publication No. US20030199016A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Yuan
; TITLE OF INVENTION: REVERSE-PHASE HPLC ASSAY FOR PLASMINOGEN ACTIVATORS
; FILE REFERENCE: PI788R1
; CURRENT APPLICATION NUMBER: US/10/443,701
; CURRENT FILING DATE: 2003-05-21
; PRIOR APPLICATION NUMBER: US/09/703,695
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/163,607
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-443-701-4

Query Match 39.2%; Score 328.5; DB 12; Length 562;
Best Local Similarity 46.3%; Pred. No. 9.3e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;
QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHGY 58
DB 77 QCHSVPVKSCSEPRCFNGGTCQALYFSDP-VQCQPEGFAGKCCIEDTRATCYEDQGISY 135
QY 59 RKGASTDTMGRCPLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVG 118
DB 136 RGTWSTAESGAECTNWNSSALAKPYSGRRPDARLGLGNHNYCRNPDNRDRSKPWCYVFK 195
QY 119 LKPLVQECMVHDCADG 134
DB 196 GKYSEFCSTPACSEG 211

RESULT 27
US-10-193-656-8
; Sequence 8, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard

; APPLICANT: LI, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/LJ577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00750
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: (1)..(562)
US-10-193-656-8

Query Match 39.2%; Score 328.5; DB 15; Length 562;
Best Local Similarity 46.3%; Pred. No. 9.3e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;
QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHGY 58
DB 77 QCHSVPVKSCSEPRCFNGGTCQALYFSDP-VQCQPEGFAGKCCIEDTRATCYEDQGISY 135
QY 59 RKGASTDTMGRCPLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVG 118
DB 136 RGTWSTAESGAECTNWNSSALAKPYSGRRPDARLGLGNHNYCRNPDNRDRSKPWCYVFK 195
QY 119 LKPLVQECMVHDCADG 134
DB 196 GKYSEFCSTPACSEG 211

RESULT 28
US-10-349-543-5
; Sequence 5, Application US/10349543
; Publication No. US20030166514A1
; GENERAL INFORMATION:
; APPLICANT: Jones, Terence R.
; Haney, David N.
; Varga, Janos
; TITLE OF INVENTION: CYCLIC PEPTIDES THAT BIND TO
; UROKINASE-TYPE PLASMINOGEN ACTIVATOR RECEPTOR
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RAUER, FISHMAN & GRAUER
; STREET: 1233 20TH STREET NW, SUITE 501
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/349,543
; FILING DATE: 22-Jan-2003
; CLASSIFICATION: <Unknown>
; PUBLICATION DATA:
; APPLICATION NUMBER: US/09/285,783
; FILING DATE: 05-Apr-1999
; ATTORNEY/AGENT INFORMATION:
; NAME: LIVNAT, SHMUEL
; REGISTRATION NUMBER: 33,949

REFERENCE/DOCKET NUMBER: ANG-001/DIV (80144-0007)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 955-8787
TELEFAX: (202) 955-3751
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Disulfide-bond
LOCATION: group(11..19, 13..31, 33..42)
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-349-543-5

Query Match 35.0%; Score 293; DB 12; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.4e-22;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCEIDKSKT 49
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCEIDKSKT 49

RESULT 29

US-09-880-503-7
Sequence 7, Application US/09880503
Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: CINES, Douglas B
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
TITLE OF INVENTION: TISSUE CONTRACTABILITY
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880,503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 323
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-7

Query Match 34.4%; Score 288; DB 10; Length 323;
Best Local Similarity 100.0%; Pred. No. 6e-21;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCEIDKSK 48
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCEIDKSK 48

RESULT 30

US-09-880-503-2
Sequence 2, Application US/09880503
Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: CINES, Douglas B
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
TITLE OF INVENTION: TISSUE CONTRACTABILITY
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880,503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2

LENGTH: 47
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-2
Query Match 33.8%; Score 283; DB 10; Length 47;
Best Local Similarity 100.0%; Pred. No. 2.4e-21;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCEIDKS 47
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFKGGQHCEIDKS 47
Search completed: December 3, 2003, 15:05:46
Job time : 15.681 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:58 ; Search time 7.39655 Seconds
(without alignments)
1859.261 Million cell updates/sec

Title: US-09-880-503-8

Perfect score: 837

Sequence: 1 SNELHQVPSNCCLNGGTCV.....QECMVHCADGKPKSPPE 143

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168692 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:*

1: Piri:*

2: Piri:*

3: Piri:*

4: Piri:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	837	100.0	431	1 UKHU	u-plasminogen acti
2	757.5	90.5	433	1 UKBAY	u-plasminogen acti
3	625.5	74.7	432	1 UKPG	u-plasminogen acti
4	605	72.3	433	1 JN0560	u-plasminogen acti
5	592	70.7	432	1 S18932	u-plasminogen acti
6	569	68.0	433	1 UKMS	u-plasminogen acti
7	335.5	40.1	477	1 A34369	t-plasminogen acti
8	335.5	40.1	477	1 J30598	t-plasminogen acti
9	334.5	40.0	434	1 A35005	u-plasminogen acti
10	328.5	39.2	291	2 J38098	t-plasminogen acti
11	328.5	39.2	431	2 J30599	t-plasminogen acti
12	328.5	39.2	562	1 UKHUT	t-plasminogen acti
13	325	38.8	559	1 A35029	t-plasminogen acti
14	315.5	37.7	559	1 A29941	t-plasminogen acti
15	310.5	37.1	477	2 J30597	t-plasminogen acti
16	268	32.0	603	2 S28941	coagulation factor
17	265	31.7	598	2 J30598	plasma hyaluronan-
18	263.5	31.5	560	1 J46588	hepatocyte growth
19	262	31.3	655	1 A46688	hepatocyte growth
20	233	27.8	615	1 KFHU12	coagulation factor
21	223	26.6	394	2 J30600	t-plasminogen acti
22	217	25.9	593	2 S45281	coagulation factor
23	169	20.2	685	1 A48289	neutrophilic recep.
24	163	19.5	460	2 B61545	plasmin (EC 3.4.21
25	162.5	19.4	4548	1 S00557	apoprotein(a) (EC
26	161	19.2	1420	2 A32869	apoprotein(a)
27	160	19.1	123	2 C61545	plasmin (EC 3.4.21
28	159.5	19.1	810	2 B30848	plasmin (EC 3.4.21
29	159	19.0	120	2 E61545	plasmin (EC 3.4.21

30	159	19.0	812	1 PLBO	plasmin (EC 3.4.21
31	158.5	18.9	937	2 A45082	neutrophilic recep
32	156.5	18.7	810	1 PLHU	plasmin (EC 3.4.21
33	155	18.5	790	1 PLPG	plasmin (EC 3.4.21
34	155	18.5	812	1 PLMS	plasmin (EC 3.4.21
35	154	18.4	89	2 A60140	plasmin (EC 3.4.21
36	153.5	18.3	810	2 I46260	plasmin (EC 3.4.21
37	152.5	18.2	169	2 A40522	plasmin (EC 3.4.21
38	150	17.9	711	1 A47136	macrophage-stimula
39	149	17.8	455	2 A61545	plasmin (EC 3.4.21
40	148	17.7	943	2 B45082	neutrophilic recep
41	147.5	17.6	716	1 A40332	macrophage-stimula
42	147	17.6	2869	2 T18518	apolipoprotein(a)
43	146.5	17.5	806	2 T18840	hypothetical prote
44	145.5	17.4	716	1 JCS061	macrophage-stimula
45	144.5	17.3	728	1 JH0579	hepatocyte growth

ALIGNMENTS

RESULT 1

UKHU

u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N:Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminog
N:Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen a
in form
C:Species: Homo sapiens (man)

C>Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text change 15-Sep-2000
C:Accession: A00931; #sequence: J0102; A37561; A37562; A37563; A37564; A
R:Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blasi, F.
Nucleic Acids Res. 13, 2759-2771, 1985
A>Title: The human urokinase-plasminogen activator gene and its promoter.
A:Reference number: A00931; MUID:85215647; PMID:2987867

A:Accession: A00931
A:Molecule type: DNA
A:Residues: 1-431 <RIC>
A:Cross-references: GB:X02419; NID:g37601; PIDN:CAA26268.1; PID:g1834524
A>Note: the authors translated the codon AIG for residue 214 as Ile
R:Nagamamine, Y.; Pearson, D.; Grattan, M.
Biochem. Biophys. Res. Commun. 132, 563-569, 1985
A>Title: Exon-intron boundary-sliding in the generation of two mRNAs coding for porcin
A:Reference number: I52209; MUID:86050639; PMID:3933505
A:Accession: I52209
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 145-161 <NAG1>
A:Cross-references: GB:K03027; NID:g340174; PIDN:AAA61257.1; PID:g340175
R:Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama
Gene 36, 183-188, 1985
A>Title: Molecular cloning of cDNA coding for human preprourokinase.
A:Reference number: J70102; MUID:86058954; PMID:2415429
A:Accession: J70102

A:Molecule type: mRNA
A:Residues: 1-213, 'I', 215-431 <NAG2>
A:Cross-references: GB:K03226; NID:g340155; PIDN:AA97138.1; PID:g340158; GB:D00244; N
R:Verde, P.; Stopelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984
A>Title: Identification and primary sequence of an unspliced human urokinase poly(A) +
A:Reference number: A37561; MUID:84272706; PMID:6589620
A:Accession: A37561

A:Molecule type: mRNA
A:Residues: 66-431 <VER>
A:Cross-references: GB:D00244; NID:g220138
R:Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Else
DNA 4, 139-146, 1985
A>Title: Molecular cloning, sequencing, and expression in Escherichia coli of human pr
A:Reference number: I38102; MUID:85203359; PMID:3888571
A:Accession: I38102

A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>
A:Cross-references: EMBL:X02760; NID:g35297; PIDN:CAA26535.1; PID:g35298

C;Keywords: fibrinolysis, glycoprotein, heterodimer, hydrolase, kringle, serine protease

F;1-20/Domain: signal sequence #status predicted <Sig>
F;21-431/Product: urokinase-type plasminogen activator, single chain form #status predicted <Sig>
F;21-177/Product: urokinase-type plasminogen activator chain A #status experimental <M>
F;31-62/Domain: EGF homology <EGF>
F;70-151/Domain: kringle homology <KRG>
F;156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental <A1>
F;179-431/Product: urokinase-type plasminogen activator chain B #status experimental
F;179-419/Domain: trypsin homology <TRY>
F;31-39,33-51,53-62,70-151,91-133,122-146,168-299,209-225,217-288,313-382,345-361,372-393,431-433/Domain: trypsin homology <TRY>
F;38/Binding site: carboxylate (Thr) (covalent) #status predicted
F;178-179/Cleavage site: Lys-Ile (plasmin) #status experimental
F;224,275,376/Active site: His, Asp, Ser #status experimental
F;322/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 100.0%; Score 837; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 9.3e-68;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHVPNSCDCLNGTCTVSNKYFNSIHWNCNPKPKGGQHCIEDKSKTCYEGNGHFFYRG 60
DB 21 SNELHVPNSCDCLNGTCTVSNKYFNSIHWNCNPKPKGGQHCIEDKSKTCYEGNGHFFYRG 80
QY 61 KASTDTMGPRCLPWNSATVLQQTIVHAHRSDALQLGLGKHNYCRPNDRRRPWCYQVGLK 120
DB 81 KASTDTMGPRCLPWNSATVLQQTIVHAHRSDALQLGLGKHNYCRPNDRRRPWCYQVGLK 140
QY 121 PLVQECMWHDCADGKKPSSPPEE 143
DB 141 PLVQECMWHDCADGKKPSSPPEE 163

RESULT 2
UKBAY

u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
C;Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
C;Accession: S14687; S08651
R;Au, Y.P.T.; Wang, T.W.; Clowes, A.W.
Nucleic Acids Res. 18, 3411, 1990
A;Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen activator
A;Reference number: S14687; MUID:90287734; PMID:2113276
A;Accession: S14687
A;Molecule type: mRNA
A;Residues: 1-433 <AUY>
A;Cross-references: EMBL:X51935; NID:G38130; PIDN:CAA36200.1; PID:G38131
C;Superfamily: urokinase-type plasminogen activator, EGF homology; kringle homology; t
C;Keywords: glycoprotein, heterodimer, hydrolase, kringle, serine proteinase
F;1-20/Domain: signal sequence #status predicted <Sig>
F;21-176/Product: plasminogen activator chain A #status predicted <ACH>
F;30-61/Domain: EGF homology <EGF>
F;69-150/Domain: kringle homology <KRG>
F;178-433/Product: plasminogen activator chain B #status predicted <BCH>
F;178-421/Domain: trypsin homology <TRY>
F;167-298,208-224,216-287,315-384,347-363,374-402/Bisulfide bonds: #status predicted
F;223,274,378/Active site: His, Asp, Ser #status predicted
F;324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 90.5%; Score 757.5; DB 1; Length 433;
Best Local Similarity 92.3%; Pred. No. 1.3e-60;
Matches 132; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 SNELHVPNSCDCLNGTCTVSNKYFNSIHWNCNPKPKGGQHCIEDKSKTCYEGNGHFFYRG 60
DB 21 SRRL-QVPDCGCLNGTCTVSNKYFNSIHWNCNPKPKGGQHCIEDKSKTCYEGNGHFFYRG 79
QY 61 KASTDTMGPRCLPWNSATVLQQTIVHAHRSDALQLGLGKHNYCRPNDRRRPWCYQVGLK 120
DB 80 KASTDTMGPRCLPWNSATVLQQTIVHAHRSDALQLGLGKHNYCRPNDRRRPWCYQVGLK 139
QY 121 PLVQECMWHDCADGKKPSSPPEE 143
DB 140 ORVQECMWHDCADGKKPSSPPEE 162

```
RESULT 3
UNP
u-plasminogen activator (EC 3.4.21.73) precursor - pig
N/Alternate names: uPA
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
C:Accession: A00932
R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
Nucleic Acids Res. 12, 9525-9541, 1984
A>Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A:Reference number: A00932; MUID:85087954; PMID:6096932
A:Accession: A00932
A:Molecule type: DNA
A:Residues: 1-240, 'H', 242-442 <NAG1>
A:Experimental source: kidney cell line LLC-PK1
R:Nagamine, Y.
submitted to the Protein Sequence Database, December 1986
A:Reference number: A37566
A:Contents: annotation; correction to residue 241
C:Genetics:
A:Introns: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:33-64/Domain: EGF homology <EGF>
F:72-153/Domain: kringle homology <KRG>
F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:190-430/Domain: trypsin homology <TRY>
F:152/Binding site: carboxylate (Asn) (covalent) #status predicted
F:179-310, 220-236, 228-299, 324-393, 356-372, 383-411/Disulfide bonds: #status predicted
F:235, 286, 387/Active site: His, Asp, Ser #status predicted

Query Match 74.7%; Score 625.5; DB 1; Length 442;
Best Local Similarity 73.4%; Pred. No. 9.1e-49;
Matches 113; Conservative 11; Mismatches 19; Indels 11; Gaps 2;

QY 1 SNELHGV--PSNCDLNGGTCVSNKYFNSHTWNCNPKKFGGCHCEIDKSKTCYEGNGHFY 58
DB 21 SHELHGESASNCGLNGKCVSYFYFNICQRCSPKPKQGGHCEIDTSQTCFEGNGHSY 80

QY 59 RGKASTDTMGRCPLPWSATVLQOQTYHAHRSALQGLGKHNYCRNPNRRPWCYQVQG 118
DB 81 RGKANTDTGRCPLPWSATVLNTYHAHRPDALQGLGKHNYCRNPNQRPWCYQVQG 140

QY 119 LKPLVQECMVHDCADGKKPSPPEE 143
DB 141 LKQLVQECMVPCNCGSGESHRPAYDGKNPFTPEK 174

RESULT 4
UN0560
u-plasminogen activator (EC 3.4.21.73) precursor - bovine
N/Alternate names: uPA
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C:Accession: JN0560
R:Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
Gene 125, 177-183, 1993
A>Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and induc
A:Reference number: JN0560; MUID:93216119; PMID:8385052
A:Accession: JN0560
A:Molecule type: mRNA
A:Residues: 1-433 <KRA>
A:Cross-references: GB:L03546; NID:q163800; PIDN:AAA51419.1; PID:q163801
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-179/Product: urokinase-type plasminogen activator chain A #status predicted <MA1>
F:21-179/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:33-64/Domain: EGF homology <EGF>
```

```
F:72-153/Domain: kringle homology <KRG>
F:181-433/Product: plasminogen activator chain B #status predicted <MA2>
F:181-421/Domain: trypsin homology <TRY>
F:170-301, 211-227, 219-290, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
F:226, 277, 378/Active site: His, Asp, Ser #status predicted

Query Match 72.3%; Score 605; DB 1; Length 433;
Best Local Similarity 72.4%; Pred. No. 6.2e-47;
Matches 105; Conservative 14; Mismatches 24; Indels 2; Gaps 1;

QY 1 SNELHGV--PSNCDLNGGTCVSNKYFNSHTWNCNPKKFGGCHCEIDKSKTCYEGNGHFY 58
DB 21 SNEVHKGESGNCGLNGKCVTYFYFNICQRCSPKPKQGGHCEIDTSQTCYEGNGHSY 80

QY 59 RGKASTDTMGRCPLPWSATVLQOQTYHAHRSALQGLGKHNYCRNPNRRPWCYQVQG 118
DB 81 RGKANDLDSGRPCLAWDSPVLLKMYHAHRSDAIQLGLGKHNYCRNPNQRRPWCYQVIG 140

QY 119 LKPLVQECMVHDCADGKKPSPPEE 143
DB 141 LKQVQCMVQDCSVGKSPSPREK 165
```

```
RESULT 5
S18932
u-plasminogen activator (EC 3.4.21.73) precursor - rat
N/Alternate names: plasminogen activator, urokinase-type, urinary plasminogen activator;
C:Species: Rattus norvegicus (Norway rat)
C:Date: 18-Oct-1989 #sequence_revision 10-Feb-1995 #text_change 18-Jun-1999
C:Accession: S24604; I60186; I53472; S18932
R:Rabbani, S.A.
submitted to the EMBL Data Library, April 1992
A:Reference number: S24604
A:Accession: S24604
A:Molecule type: mRNA
A:Residues: 1-15, 'H', 17-23, 'G', 25-331, 'N', 333-432 <RAB>
A:Cross-references: EMBL:X65651; NID:g57456; PIDN:CAA46601.1; PID:g57457
A:Experimental source: tissue kidney
R:Henderson, B.R.; Tansy, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.
Cancer Res. 52, 2489-2496, 1992
A>Title: Transcriptional and posttranscriptional activation of urokinase plasminogen a
A:Reference number: I60186; MUID:92233409; PMID:1568219
A:Accession: I60186
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-432 <RES>
A:Cross-references: EMBL:X63434; NID:g57465; PIDN:CAA45028.1; PID:g57466
A:Experimental source: strain Fischer 344; tissue mammary
R:Ragno, P.; Cassano, S.; Degen, J.; Kessler, C.; Blasi, F.; Rossi, G.
FEBS Lett. 306, 193-198, 1992
A>Title: The receptor for the plasminogen activator of urokinase type is up-regulated :
A:Reference number: I53472; MUID:92339549; PMID:1321734
A:Accession: I53472
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 31-62 <RE2>
A:Cross-references: EMBL:X66907; NID:g396200; PIDN:CAA47356.1; PID:g938279
C:Genetics:
A:Gene: uPA
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:31-62/Domain: EGF homology <EGF>
F:70-151/Domain: kringle homology <KRG>
F:179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:179-420/Domain: trypsin homology <TRY>
F:168-300, 210-226, 218-289, 314-383, 346-362, 373-401/Disulfide bonds: #status predicted
F:225, 276, 377/Active site: His, Asp, Ser #status predicted

Query Match 70.7%; Score 592; DB 1; Length 432;
Best Local Similarity 75.6%; Pred. No. 9.1e-46;
Matches 102; Conservative 10; Mismatches 23; Indels 0; Gaps 0;
```

```

QY 9 SNCDCLNGGTCVSNKYFNSIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRKGASTDTMG 68
DB 29 SNGCGQNGGVCVSKYKFSIRRCSPKPKGHECEIDTSKTCYHGNGQSYRGKANTDTKG 88
QY 69 RPLCPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLKPLVQECMV 128
DB 89 RPLCLANNAVPAVLQTYNAHRSALSLGLGKHNYCRNPNRRPWCYVQVGLKQFVQECMV 148
QY 129 HDCADGKKSPSPPEE 143
DB 149 QDCSLSKSPSSVDQ 163

RESULT 6
UKMS
u-plasminogen activator (EC 3.4.21.73) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
C:Accession: A29420; A24615
R:Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
Biochemistry 26, 8270-8279, 1987
A:Title: The murine urokinase-type plasminogen activator gene.
A:Reference number: A29420; MUID:88163489; PMID:2831940
A:Accession: A29420
A:Molecule type: DNA
A:Residues: 1-433 <DEG>
A:Cross-references: GB:M17922; NID:G202296; PIDN:AAA40539.1; PID:G202297
R:Belin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Kocher,
Eur. J. Biochem. 148, 225-232, 1985
A:Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase
A:Reference number: A24615; MUID:85179474; PMID:2985383
A:Accession: A24615
A:Molecule type: mRNA
A:Residues: 1-433 <BEL>
A:Cross-references: GB:X02389; NID:G55127; PIDN:CAA26231.1; PID:G55128
C:Genetics:
A:Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:32-63/Domain: EGF homology <EGF>
F:71-152/Domain: kringle homology <KRG>
F:180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:180-421/Domain: trypsin homology <TRY>
F:169-301/211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
F:226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 68.0%; Score 569; DB 1; Length 433;
Best Local Similarity 70.4%; Pred. No. 1.1e-43;
Matches 95; Conservative 15; Mismatches 25; Indels 0; Gaps 0;

QY 9 SNCDCLNGGTCVSNKYFNSIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRKGASTDTMG 68
DB 30 SNGCGQNGGVCVSKYKFSIRRCSPKPKGHECEIDASKTCYHGNGDSYRGKANTDTKG 89
QY 69 RPLCPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLKPLVQECMV 128
DB 90 RPLCLANNAVPAVLQTYNAHRSALSLGLGKHNYCRNPNRRPWCYVQVGLKQFVQECMV 149
QY 129 HDCADGKKSPSPPEE 143
DB 150 HDCSLSKSPSSVDQ 164

RESULT 7
A34369
t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
C:Species: Megaderma lyra
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A34369
R:Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacob

```

```

J. Biol. Chem. 264, 17947-17952, 1989
A:Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmi
A:Reference number: A34369; MUID:90038667; PMID:2509450
A:Accession: A34369
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-477 <GAR>
A:Cross-references: GB:J05082; NID:G166080; PIDN:AAA31596.1; PID:G166081
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <1FA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: kringle homology <KRG>
F:226-471/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-
F:272,321,428/Active site: His, Asp, Ser #status predicted
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 40.1%; Score 335.5; DB 2; Length 477;
Best Local Similarity 46.3%; Pred. No. 1e-22;
Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;

QY 3 ELHQVP----SNCDCLNGGTCVSNKYFNSIHWCNCPKFGGQHCIEDKSKTCYEGNGHYF 58
DB 78 QCHTVPVKSCSELRCFNGGTCWQAASFDF--VCQCPKGYTGKQCEVDTHTATCYKQDQGVY 136
QY 59 RGKASTDMGRPCIPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVQ 118
DB 137 RGTWSTESGAGQCIWNSNLTRTYNGRRSDAITLGLGHNYCRNPNRRPWCYVQVQ 196
QY 119 LKPLVQECMVHDCA 132
DB 197 SKFILEFCVPPVCS 210

RESULT 8
J50598
t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: J50598
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; D
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat D
A:Reference number: J50597; MUID:92039036; PMID:1937019
A:Accession: J50598
A:Molecule type: mRNA
A:Residues: 1-477 <KRA>
A:Cross-references: GB:M63988; NID:G166074; PIDN:AAA31593.1; PID:G166075
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <1FA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: kringle homology <KRG>
F:226-471/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-
F:185,398/Binding site: carboxylate (Asn) (covalent) #status predicted
F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
F:272,321,428/Active site: His, Asp, Ser #status predicted
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 40.1%; Score 335.5; DB 2; Length 477;
Best Local Similarity 46.3%; Pred. No. 1e-22;
Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;

QY 3 ELHQVP----SNCDCLNGGTCVSNKYFNSIHWCNCPKFGGQHCIEDKSKTCYEGNGHYF 58
DB 78 QCHTVPVKSCSELRCFNGGTCWQAASFDF--VCQCPKGYTGKQCEVDTHTATCYKQDQGVY 136

```

```
QY 59 RGKASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDRRPPWCYVQVG 118
Db 137 RGTWSTESGAQCINN SLLTRRTYNGRRSDAITLGLGNHNYCRNPDRNPKWCYVKA 196
QY 119 LKPLVQECMVHDC A 132
Db 197 SKFILEFCSPVPCS 210

RESULT 9
A35005
u-plasminogen activator (EC 3.4.21.73) precursor - chicken
N:Alternate names: uPA
C:Species: Gallus gallus (chicken)
C:Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 16-Jul-1999
C:Accession: A35005
J:Leslie, N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.
R. Biol. Chem. 265, 1339-1344, 1990
A:Title: The chicken urokinase-type plasminogen activator gene.
A:Reference number: A35005; MUID:90110185; PMID:2295632
A:Accession: A35005
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-434 <LES>
A:Cross-references: GB:J05187; NID:G212858; PIDN:AAA49131.1; PID:G212859
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:40-71/Domain: EGF homology <EGF>
F:79-158/Domain: kringle homology <KRG>
F:173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:173-416/Domain: trypsin homology <TRY>
F:162-296,202-218,210-285,310-379,342-358,369-397/Disulfide bonds: #status predicted
F:217,272,373/Active site: His, Asp, Ser #status predicted

Query Match 40.0%; Score 334.5; DB 1; Length 434;
Best Local Similarity 54.8%; Pred. No. 1.2e-22;
Matches 58; Conservative 14; Mismatches 30; Indels 5; Gaps 2;

QY 11 CDCLNGGTCVSNKYFSNIHWCNPKKFGQGHCEIDKSKTCYEGNGHFGYRGKASTDTMGRP 70
Db 40 CQCLNGGTCITVRFQIKRCLCPGYGLHCEIDTNSICSGNGEDYRGMAEDP----G 95
QY 71 CLPWN SATVLQ-QTYHAHRS DALQLGLGKHNYCRNPDRRPPWCYVQ 116
Db 96 CLYWDHPSVIRMGDYHADLKNALQLGLGKHNYCRNPNGSRPWCYTK 142

RESULT 10
I38098
t-plasminogen activator precursor, inactive endothelial splice form - human
N:Alternate names: tissue plasminogen activator
C:Species: Homo sapiens (man)
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
C:Accession: I38098; S01678
R:Siebert, P.D.; Fong, K.
Nucleic Acids Res. 18, 1086, 1990
A:Title: Variant tissue type plasminogen activator (PLAT) cDNA obtained from human endot
A:Reference number: I38098; MUID:90192128; PMID:1969145
A:Accession: I38098
A:Status: translated
A:Molecule type: mRNA
A:Residues: 1-291 <SRI>
A:Cross-references: EMBL:X13057; NID:G35282; PIDN:CAA31489.1; PID:G35283
C:Comment: For the main splice form, see PIR:UKHUT. This form probably does not have pro
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
```

```
C:Keywords: alternative splicing; fibrinolysis; glycoprotein; kringle
F:1-33/Domain: signal sequence #status predicted <SIG>
F:24-32/Domain: propeptide #status predicted <PRO>
F:33-291/Product: t-plasminogen activator, inactive endothelial splice form #status pr
F:41-78/Domain: fibronectin type I repeat homology <1FA>
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology <KR1>
F:215-291/Domain: kringle homology #status atypical <KR2>
F:41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203/Disulfide bonds: #status pr

Query Match 39.2%; Score 328.5; DB 2; Length 291;
Best Local Similarity 46.3%; Pred. No. 2.9e-22;
Matches 63; Conservative 13; Mismatches 55; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNPKKFGQGHCEIDKSKTCYEGNGHFGY 58
Db 77 QCHSVPVKSCSEPRCFNGGTCQALYFSDF-VCCQPEGFAGKCEIDTRATCYEDQGILSY 135
QY 59 RGKASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDRRPPWCYVQVG 118
Db 136 RGTWSTAESGAECTNWNSSALANAYSGRRPDALRLGLGNHNYCRNPDRDSKPCWCYVKA 195
QY 119 LKPLVQECMVHDCADG 134
Db 196 GYSESEFCSTPACSEG 211

RESULT 11
J50599
t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: J50599
R:Kraetschmar, J.; Haendler, B.; Langer, G.; Boisdol, W.; Bringmann, P.; Alagon, A.; De
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat D
A:Reference number: J50597; MUID:92039036; PMID:1937019
A:Accession: J50599
A:Molecule type: mRNA
A:Residues: 1-431 <KRA>
A:Cross-references: GB:M6399; NID:G166076; PIDN:AAA31594.1; PID:G166077
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-431/Product: plasminogen activator beta #status predicted <PLA>
F:41-74/Domain: EGF homology <EGF>
F:82-163/Domain: kringle homology <KRG>
F:180-425/Domain: trypsin homology <TRY>
F:41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disulfide b
F:139,352/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:179-180/Cleavage site: His-Ser (plasmin) #status predicted
F:226,275,382/Active site: His, Asp, Ser #status predicted
F:345-361,378-406/Disulfide bonds: #status predicted

Query Match 39.2%; Score 328.5; DB 2; Length 431;
Best Local Similarity 47.6%; Pred. No. 4e-22;
Matches 59; Conservative 16; Mismatches 48; Indels 1; Gaps 1;

QY 9 SNCDCLNGGTCVSNKYFSNIHWCNPKKFGQGHCEIDKSKTCYEGNGHFGYRGKASTDTMG 68
Db 42 SELRCFNGGTCQAAASFDF-VCCQPKGYTGKQCEVDTHATCYKDQGVYRGWTSTSESG 100
QY 69 RPTCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDRRPPWCYVQVGLKPLVQECMV 128
Db 101 AQCINWNSNLLTRTYNGRRSDAITLGLGNHNYCRNPDRNPKWCYVKAIFLEFCV 160
QY 129 HDCA 132
Db 161 PVCS 164
```

RESULT 12

t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
 N;Alternate names: t-PA; tissue plasminogen activator
 C;Species: Homo sapiens (man)
 C;Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text change 08-Dec-2000
 C;Accession: A94004; A23529; J05562; A93293; S02125; A91343; A93951; A91322; A54645; 160
 R;Ny, T.; Elgh, F.; Lund, B.
 Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
 A;Title: The structure of the human tissue-type plasminogen activator gene: correlation
 A;Reference number: A94004; MUID:84298137; PMID:6089198
 A;Accession: A94004
 A;Molecule type: DNA
 A;Residues: 1-562 <NVT>
 A;Cross-references: GB:100141
 A;Note: the codon given for residue 93 (ACC) is inconsistent with the authors' translation
 R;Friedman, S.J.; Rajput, B.; Reich, E.
 J. Biol. Chem. 261, 6972-6985, 1986
 A;Title: The human tissue plasminogen activator gene.
 A;Reference number: A23529; MUID:86196143; PMID:3009482
 A;Accession: A23529
 A;Molecule type: DNA
 A;Residues: 1-562 <DEG>
 A;Cross-references: GB:X03021; NID:g339817; PIDN:AAA98809.1; PID:g339819
 R;Iwagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.
 Agric. Biol. Chem. 55, 1225-1232, 1991
 A;Title: Purification and characterization of tissue plasminogen activator secreted by H
 A;Reference number: J05562; MUID:91291340; PMID:1368681
 A;Accession: J05562
 A;Molecule type: mRNA
 A;Residues: 31-562 <ITA>
 A;Cross-references: DBJ:D01096; NID:g220128; PIDN:BA00881.1; PID:g441174
 A;Experimental source: embryonic lung fibroblast IMR-90 cells
 A;Note: part of this sequence, including the amino end of the mature protein, was confir
 R;Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vohar, G.A.; Ward, C.A.; Bennett
 Nature 301, 214-221, 1993
 A;Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche
 A;Reference number: A93293; MUID:83115262; PMID:6337343
 A;Accession: A93293
 A;Molecule type: mRNA
 A;Residues: 1-562 <PEN>
 A;Cross-references: GB:100141
 A;Experimental source: melanoma cells
 R;Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.
 Nucleic Acids Res. 16, 5695, 1988
 A;Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe
 A;Reference number: S02125; MUID:88262579; PMID:3133640
 A;Accession: S02125
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-562 <SAS>
 A;Cross-references: EMBL:X07393; NID:g37243; PIDN:CAA30302.1; PID:g37244
 A;Experimental source: fetal lung cells
 R;Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma
 FEBS Lett. 189, 145-149, 1985
 A;Title: Expression in *Escherichia coli* of finger-domain lacking tissue-type plasminogen
 A;Reference number: A91343; MUID:85285620; PMID:3896853
 A;Accession: A91343
 A;Molecule type: mRNA
 A;Residues: 1-38, 'G', 86-433, 'E', 435-562 <KAG>
 A;Experimental source: Detroit 562 cells; ATCC 138
 R;Edlund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
 Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
 A;Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activ
 A;Reference number: A93951; MUID:83169656; PMID:6572897
 A;Accession: A93951
 A;Molecule type: mRNA
 A;Residues: 251-358 <EDI>
 A;Experimental source: melanoma cells
 R;Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.
 Biochemistry 23, 3701-3707, 1984
 A;Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am
 differences.

A;Reference number: A90488; MUID:85000468; PMID:6433976
 A;Contents: annotation; melanoma cells, partial sequence of residues 36-562, active an
 R;Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.
 FEBS Lett. 168, 29-32, 1984
 A;Title: Differences between uterine and melanoma forms of tissue plasminogen activator
 A;Reference number: A91322; MUID:84159956; PMID:6538514
 A;Accession: A91322
 A;Molecule type: protein
 A;Residues: 33-45;311-320 <POH>
 A;Experimental source: uterus
 A;Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
 R;Van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.
 J. Biol. Chem. 261, 14214-14218, 1986
 A;Reference number: A37567; MUID:87033611; PMID:3021732
 A;Contents: annotation; fibrin binding site
 R;Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Eng
 EMBO J. 5, 3525-3530, 1986
 A;Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen
 A;Reference number: A37568; MUID:87161761; PMID:3030730
 A;Contents: annotation; fibrin binding site
 R;Dodd, I.; Nunn, B.; Robinson, J.H.
 Thromb. Haemost. 59, 523-528, 1988
 A;Title: Isolation, identification and pharmacokinetic properties of human tissue-type
 A;Reference number: A50902; MUID:89044681; PMID:3142086
 A;Contents: annotation; novel forms of expressed recombinant t-PA
 R;Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Opendakker, G.
 Mol. Biol. Med. 3, 279-292, 1986
 A;Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its ex
 A;Reference number: A54645; MUID:86284200; PMID:3090401
 A;Accession: A54645
 A;Molecule type: mRNA
 A;Residues: 1-562 <HAR>
 A;Cross-references: GB:M15518; NID:g130031; PIDN:AAA60111.1; PID:g130032
 A;Note: parts of this sequence were confirmed by peptide sequencing
 R;Reddy, V.B.; Garramone, A.J.; Sasaki, H.; Wei, C.
 DNA 6, 461-472, 1987
 A;Title: Expression of human uterine tissue-type plasminogen activator in mouse cells
 A;Reference number: I60110; MUID:88054470; PMID:2824147
 A;Accession: I60110
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-562 <RES>
 A;Cross-references: GB:M18192; NID:g340176; PIDN:AAA36800.1; PID:g340177
 R;Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.
 J. Biol. Chem. 260, 11223-11230, 1985
 A;Title: Isolation and characterization of the human tissue-type plasminogen activator
 A;Reference number: I55232; MUID:85289338; PMID:3161893
 A;Accession: I55232
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-36 <RE2>
 A;Cross-references: GB:M1890; NID:g339837; PIDN:AAA61213.1; PID:g339839
 C;Comment: Cleavage by plasmin or trypsin produces two chains held together by a single
 C;Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-val bond.
 C;Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat
 C;Genetics:
 A;Gene: GDB:PLAT
 A;Cross-references: GDB:119496; OMIM:173370
 A;Map position: 8p12-8p12
 A;Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 5
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
 F;1-23/Domain: signal sequence #status predicted <SIG>
 F;24-32/Domain: propeptide #status predicted <PRO>
 F;33-562/Product: t-plasminogen activator #status experimental <VAT>
 F;33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
 F;41-78/Domain: fibronectin type I repeat homology <IFI>
 F;86-119/Domain: EGF homology <EGF>
 F;127-208/Domain: kringle homology <KR1>
 F;215-296/Domain: kringle homology <KR2>
 F;311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
 F;311-556/Domain: trypsin homology <TRY>
 F;41-71,69-78,86-97,91-106,110-119,121-208,148-190,179-203,215-296;236-278,267-291,299

C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C/Accession: J50597
R/Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don
Gene 105, 229-237, 1991
A/Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A/Reference number: J50597; MUID:92039036; PMID:1937019
A/Accession: J50597
A/Molecule type: mRNA
A/Residues: 1-477 <KRA>
A/Cross-references: GB:M63987; NID:G166070; PIDN:AAA31591.1; PID:G166071
C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-47/Product: plasminogen activator alpha-1 #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <IFA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: kringle homology <KRG>
F:226-471/Domain: trypsin homology <TRY>
F:42-72, 70-79, 98-92, 109, 111-120, 128-209, 149-191, 180-204, 214-345, 257-273, 265-334, 359-4
F:153, 398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
F:272, 321, 428/Active site: His, Asp, Ser #status predicted

Query Match 37.1%; Score 310.5; DB 2; Length 477;
Best Local Similarity 45.5%; Pred. No. 1.8e-20;
Matches 60; Conservative 14; Mismatches 53; Indels 5; Gaps 3;

QY 5 HQVPSN-CD---CLNGGTCVSNKYFNSIHWCNCPKFGQGHCEIDKSKTCYEGNGHFVRG 60
Db 80 HTVPVNSCEPCFNGFCNQAVYDF-VCCQFAGYTKRCDEVTRAFYEGQGVYRG 138

QY 61 KASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
Db 139 TWSTAESRVEICWNSSLLTRTYNGRMPDAFNLGLGNHNYCRNPNRNGAPKPCYVIKAGK 198

QY 121 PLVQECNWHDC 132
Db 199 FTSSECSVPVCS 210

RESULT 16
S28941
coagulation factor Xila (EC 3.4.21.38) - guinea pig (fragment)
N/Alternate names: Hageman factor
C/Species: Cavia porcellus (guinea pig)
C/Date: 25-Feb-1994 #sequence_revision 03-Aug-1995 #text_change 21-Jan-2000
C/Accession: S28941
R/Semba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kanbara, T.; Okabe, H.
Biochim. Biophys. Acta 1159, 113-121, 1992
A/Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage si
A/Reference number: S28941; MUID:93003367; PMID:1390917
A/Accession: S28941
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-603 <SEM>
A/Cross-references: EMBL:X68615; NID:G49578; PIDN:CAA48600.1; PID:G49579
C/Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C/Keywords: hydrolase; serine proteinase
F:46-87/Domain: fibronectin type II repeat homology <IF2>
F:134-169/Domain: fibronectin type I repeat homology <IF1>
F:177-208/Domain: EGF homology <EGF>
F:216-294/Domain: kringle homology <KRG>
F:359-597/Domain: trypsin homology <TRY>

Query Match 32.0%; Score 268; DB 2; Length 603;
Best Local Similarity 38.6%; Pred. No. 1.4e-16;
Matches 51; Conservative 21; Mismatches 52; Indels 8; Gaps 3;

QY 13 CLNGGTCVSNKYFNSIHWCNCPKFGQGHCEIDKSKTCYEGNGHFVRGKASDTMGPRCL 72
Db 182 CLNGGRCLF---VEGHLDCPCMGVYTGPCDDLTATSCYEGRGVSRGARTVTSAGAKC 238

QY 73 PWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLKPLVQECNWHDC 131
Db 239 RWAS-----EATYRNTAEQALRGLGHHTFCRPNDDNTRPWCFFVMGNRLSWCYDLAQ 294

QY 132 ADGKKPSSPPE 143
Db 295 QYFPQPTATPHD 306

RESULT 17
JC5878
plasma hyaluronan-binding protein precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 16-Jul-1999
C/Accession: JC5878
R/Hashimoto, K.; Tobe, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura, I
Biol. Pharm. Bull. 20, 1127-1130, 1997
A/Title: Cloning of the cDNA for a mouse homologue of human PHBP: A novel hyaluronan-b
A/Reference number: JC5878; MUID:98065239; PMID:9401717
A/Accession: JC5878
A/Molecule type: mRNA
A/Residues: 1-558 <HAS>
C/Comment: This protein acts as serine protease.
C/Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; try
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-311/Product: plasma hyaluronan-binding protein large chain #status predicted <MAT
F:75-106/Domain: EGF homology <EG1>
F:113-145/Domain: EGF homology <EG2>
F:152-185/Domain: EGF homology <EG3>
F:192-274/Domain: kringle homology <KRI>
F:312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <MA
F:312-548/Domain: trypsin homology <TRY>

Query Match 31.7%; Score 265; DB 2; Length 558;
Best Local Similarity 41.1%; Pred. No. 2.5e-16;
Matches 53; Conservative 19; Mismatches 51; Indels 6; Gaps 4;

QY 13 CLNGGTCVSNKYFNSIHWCNCPKFGQGHCEIDKSKTCYEGNGHFVRGKASDTMGPRCL 72
Db 157 CQGGVCSRRHRSRF-TCACPDQYKGFCEIGPD-DCYVGDGYSYRGKYSKTVNQNPCL 214

QY 73 PWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLKPLVQECNWHDC 131
Db 215 YWNSHLLQETYNFMFEDAEHTGTAEHFCRNPDDGDKHPWCFFVMGNRLSWCYDLAQ 274

QY 132 ADGKKPSSP 140
Db 275 ---PVPDTP 280

RESULT 18
JC4795
plasma hyaluronan-binding protein precursor - human
N/Alternate names: hepatocyte growth factor activator-like protein; PHBP
N/Contains: serine proteinase (man)
C/Species: Homo sapiens (man)
C/Date: 15-Oct-1995 #sequence_revision 16-Aug-1996 #text_change 19-Jul-2002
C/Accession: JC4795
R/Choi-Miura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Mazda, T.; Tomita, M.
J. Biochem. 119, 1157-1165, 1996
A/Title: Purification and characterization of a novel hyaluronan-binding protein (PHBP
r activator.
A/Reference number: JC4795; MUID:96425001; PMID:8827452
A/Accession: JC4795
A/Molecule type: mRNA
A/Residues: 1-560 <CHO>
A/Cross-references: GB:S83182; NID:G1836158; PIDN:AAB46909.1; PID:G1836159
A/Experimental source: Plasma
A/Note: parts of this sequence, including the amino ends of the mature chains, were de
C/Genetics:
A/Gene: GDB:HABP2; HABP; PHBP; HGFAL
A/Cross-references: GDB:4573962
C/Complex: a disulfide-bonded heterodimer of chains produced from the same precursor;

C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsin
 C:KeyWords: chondroitin sulfate proteoglycan; glycoprotein; hyaluronic acid; hydrolase;
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-113/Product: plasma hyaluronan-binding protein, 50K chain #status predicted <50K>
 F:77-108/Domain: EGF homology <EG1>
 F:115-147/Domain: EGF homology <EG2>
 F:154-187/Domain: EGF homology <EG3>
 F:194-276/Domain: kringle homology <KRI>
 F:314-550/Domain: trypsin homology <TRY>
 F:314-516/Product: plasma hyaluronan-binding protein, catalytic chain #status predicted
 F:340-207/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:77-88,82-97,99-108,115-120,136,138-147,154-165,159-176,178-187,194-276,215-257,246
 F:362,405,509/Active site: His, Asp, Ser #status predicted

Query Match 31.5%; Score 263.5; DB 1; Length 560;
 Best Local Similarity 41.0%; Pred. No. 3.4e-16;
 Matches 55; Conservative 18; Mismatches 56; Indels 5; Gaps 4;

QY 13 CLNGGTCVSNKVFNIHWCNPKKGGQHCETDKSKTYEGNHGPRYRKASTDTWGRPL 72
 Db 159 CQNGATCSRHKRSKF-TCACPDQKGFCEIG-SDDCYVGGDGYSGYRGKMNRTVNOHACL 216
 QY 73 PWSATVLQOTYHAHRSDALQLGLGKHNYCRNRRPWCYVQVGLKPLVQE-CMVHDC 131
 Db 217 YNSHLLQENYMPWEDAEHGEHFCRNPDADKFWCEIKVTNDKVKWEYCDVSAC 276
 QY 132 A--DGKXPSPPEE 143
 Db 277 SAQDVAYPESPT 290

RESULT 19
 A:6688
 C:Superfamily: growth factor activator (EC 3.4.21.-) precursor [validated] - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
 C:Accession: A46688
 R:Miyauchi, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
 J. Biol. Chem. 268, 10024-10028, 1993
 A:Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease
 d coagulation factor XII.
 A:Reference number: A46688; MUID:93252878; PMID:7683665
 A:Accession: A46688
 A:Molecule type: mRNA
 A:Residues: 1-655 <MIV>
 A:Cross-references: DBJ:U14012; NID:9219680; PIDN:AAA03113.1; PID:9219681
 A:Experimental source: liver (mRNA); serum (protein)
 A:Note: sequence extracted from NCBI backbone (NCBIN:131227, NCBI:131228)
 C:Genetics:
 A:Gene: GDB:HGPAC; HGFA: HGPAP
 A:Cross-references: GDB:9954514
 A:Map position: 4p16-4p16
 C:Function:
 A:Description: activates hepatocyte growth factor by specific proteolytic cleavage
 A:Pathway: tissue repair and regeneration
 C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
 C:KeyWords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase
 F:1-34/Domain: signal sequence #status predicted <SIG>
 F:108-148/Domain: fibronectin type II repeat homology <IF2>
 F:164-197/Domain: EGF homology <EG1>
 F:202-237/Domain: fibronectin type I repeat homology <IF1>
 F:245-278/Domain: EGF homology <EG2>
 F:286-367/Domain: kringle homology <KRG>
 F:373-407/Product: hepatocyte growth factor activator light chain #status experimental
 F:408-655/Product: hepatocyte growth factor activator heavy chain #status experimental
 F:408-641/Domain: trypsin homology <TRY>
 F:404,48,290,468,492,546/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:164-175,169-186,188-197,202-230,228-237,245-256,250-267,269-278,286-367,307-349,338-36
 F:447,497,538/Active site: His, Asp, Ser #status predicted

Query Match 31.3%; Score 262; DB 1; Length 655;
 Best Local Similarity 36.9%; Pred. No. 5.3e-16;

Matches 58; Conservative 12; Mismatches 57; Indels 30; Gaps 3;
 QY 5 HQVPSNCDCLNGTCTCVSNKVFNIHWCNPKKPGGQHCETDKSKTYEGNHG 56
 Db 242 HTACLSPLCLNGTCTCVSNKVFNIHWCNPKKPGGQHCETDKSKTYEGNHG 292
 QY 57 FYRGKASTDTWGRPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNRRPWCYVQ 116
 Db 293 GYRGVASTSASGLSCLAWSDDLTYQELHVDVSGAALLGLGPHAYCRNPNDRPWCYV 352
 QY 117 VGLKPLVQECMVHDC-----ADGKXPSP 140
 Db 353 KDSALSWEYCRLEACESLTVQLSPDLLATLPBPASP 389

RESULT 20
 KFHUI2
 C:Superfamily: growth factor activator (EC 3.4.21.38) precursor [validated] - human
 N:Alternate names: Hageman factor (activated)
 C:Species: Homo sapiens (man)
 C:Date: 27-Nov-1985 #sequence_revision 30-Jun-1991 #text_change 08-Dec-2000
 C:Accession: A29411; A26814; A00930; A25191; A22248; A21037
 R:Cool, D.E.; MacGillivray, R.F.A.
 J. Biol. Chem. 262, 13662-13673, 1987
 A:Title: Characterization of the human blood coagulation factor XII gene. Intron/exon
 A:Reference number: A29411; MUID:88007593; PMID:2888762
 A:Accession: A29411
 A:Molecule type: DNA
 A:Residues: 1-615 <COO>
 A:Cross-references: GB:M17466; GB:J02807; NID:gl80355; PIDN:AAB59490.1; PID:gl80357
 R:Tripos, M.; Citarella, F.; Guida, S.; Galeffi, P.; Fantoni, A.; Cortese, R.
 Nucleic Acids Res. 14, 3146, 1986
 A:Title: CDNA sequence coding for human coagulation factor XII (Hageman).
 A:Reference number: A26814; MUID:86176794; PMID:3754331
 A:Accession: A26814
 A:Molecule type: mRNA
 A:Residues: 4-615 <TRI>
 A:Cross-references: GB:M3115; NID:gl82291; PIDN:AAA70225.1; PID:gl82292
 R:Cool, D.E.; Edgell, C.J.S.; Louie, G.V.; Zoller, M.J.; Brayer, G.D.; MacGillivray, R.
 J. Biol. Chem. 260, 13666-13676, 1985
 A:Title: Characterization of human blood coagulation factor XII cDNA. Prediction of the
 A:Reference number: A00930; MUID:86033830; PMID:3877053
 A:Accession: A00930
 A:Molecule type: mRNA
 A:Residues: 14-332; S'334-615 <CO2>
 A:Cross-references: GB:M11723; NID:gl80358; PIDN:AAA51986.1; PID:gl80359
 R:Que, B.G.; Davie, E.W.
 Biochemistry 25, 1525-1528, 1986
 A:Title: Characterization of a cDNA coding for human factor XII (Hageman factor).
 A:Reference number: A25191; MUID:86216049; PMID:3011063
 A:Accession: A25191
 A:Molecule type: mRNA
 A:Residues: 146-378; G'380-615 <QUE>
 A:Cross-references: GB:M13147; NID:gl80360; PIDN:AAA70224.1; PID:gl80361
 R:McMullen, B.A.; Fujikawa, K.
 J. Biol. Chem. 260, 5328-5341, 1985
 A:Title: Amino acid sequence of the heavy chain of human alpha-factor XIIa (activated).
 A:Reference number: A22248; MUID:85182674; PMID:3886654
 A:Accession: A22248
 A:Molecule type: protein
 A:Residues: 20-379 <MCM>
 R:Fujikawa, K.; McMullen, B.A.
 J. Biol. Chem. 258, 10924-10933, 1983
 A:Title: Amino acid sequence of human beta-factor XIIa.
 A:Reference number: A21037; MUID:83291041; PMID:6604055
 A:Accession: A21037
 A:Molecule type: protein
 A:Residues: 354-362; 373-615 <FUI>
 R:Harris, R.J.; Ling, V.T.; Spellman, M.W.
 J. Biol. Chem. 267, 5102-5107, 1992
 A:Title: O-linked fucose is present in the first epidermal growth factor domain of fac
 A:Reference number: A4406; MUID:92184750; PMID:1544894
 A:Contents: annotation; carbohydrate binding site

A>Note: several genes closely linked on chromosome 6 are identical in the first coding exons of kringle repeats
 C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology
 C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-4548/Product: apolipoprotein(a) #status experimental <MAT>
 F:28-105/Domain: kringle homology <KR1>
 F:142-219/Domain: kringle homology <KR2>
 F:256-333/Domain: kringle homology <KR3>
 F:370-447/Domain: kringle homology <KR4>
 F:484-561/Domain: kringle homology <KR5>
 F:598-675/Domain: kringle homology <KR6>
 F:712-789/Domain: kringle homology <KR7>
 F:826-903/Domain: kringle homology <KR8>
 F:940-1017/Domain: kringle homology <KR9>
 F:1054-1131/Domain: kringle homology <KR10>
 F:1168-1245/Domain: kringle homology <KR11>
 F:1282-1359/Domain: kringle homology <KR12>
 F:1396-1473/Domain: kringle homology <KR13>
 F:1510-1587/Domain: kringle homology <KR14>
 F:1624-1701/Domain: kringle homology <KR15>
 F:1738-1815/Domain: kringle homology <KR16>
 F:1852-1929/Domain: kringle homology <KR17>
 F:1966-2043/Domain: kringle homology <KR18>
 F:2080-2157/Domain: kringle homology <KR19>
 F:2194-2271/Domain: kringle homology <KR20>
 F:2308-2385/Domain: kringle homology <KR21>
 F:2422-2499/Domain: kringle homology <KR22>
 F:2536-2613/Domain: kringle homology <KR23>
 F:2650-2727/Domain: kringle homology <KR24>
 F:2764-2841/Domain: kringle homology <KR25>
 F:2878-2955/Domain: kringle homology <KR26>
 F:2992-3069/Domain: kringle homology <KR27>
 F:3106-3183/Domain: kringle homology <KR28>
 F:3220-3297/Domain: kringle homology <KR29>
 F:3334-3411/Domain: kringle homology <KR30>
 F:3448-3525/Domain: kringle homology <KR31>
 F:3562-3639/Domain: kringle homology <KR32>
 F:3676-3753/Domain: kringle homology <KR33>
 F:3782-3859/Domain: kringle homology <KR34>
 F:3896-3973/Domain: kringle homology <KR35>
 F:4010-4087/Domain: kringle homology <KR36>
 F:4124-4201/Domain: kringle homology <KR37>
 F:4228-4307/Domain: kringle homology <KR38>
 F:4328-4541/Domain: trypsin homology <TRY>
 Query Match 19.4%; Score 162.5; DB 1; Length 4548;
 Best Local Similarity 33.1%; Pred. No. 2.4e-06;
 Matches 41; Conservative 12; Mismatches 46; Indels 25; Gaps 7;
 QY 27 NIHW--CN---CP-----KKFGQHCIDKSKT---CYEGNGHFYRGKASTDTMGR 69
 Db 3742 NVRREYCNLTCCPVTESVSLATSTAVSEQAPTQSPVQDCYHGDQSGYRGSTTTVGR 3801
 QY 70 PCLPWNATVLTQYTHAHRSDALQLGLGKKNYCRNPNRRPWCYVOVGLKPLV--QECM 127
 Db 3802 TCOSWSSMT---PHWHORTTEYPNGGLTRNYCRNPDAIRPWCYT---MDPSVRWEYCN 3855
 QY 128 VHDC 131
 Db 3856 LTQC 3859
 RESULT 26
 A32869
 apolipoprotein(a) (EC 3.4.21.-) - rhesus macaque (fragment)
 C:Species: Macaca mulatta (rhesus macaque)
 C>Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 22-Jun-1999
 C:Accession: A32869; A30848
 R:Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.
 J. Biol. Chem. 264, 5957-5965, 1989
 A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
 A:Reference number: A32869; MUID:89174660; PMID:2925643

A:Accession: A32869
 A:Molecule type: mRNA
 A:Residues: 1-1420 <TM>
 C:Cross-references: GB:J04635; NID:g342072; PIDN:AAA36833.1; PID:g342073
 C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology
 C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
 F:50-127/Domain: kringle homology <KR1>
 F:164-241/Domain: kringle homology <KR2>
 F:278-355/Domain: kringle homology <KR3>
 F:392-469/Domain: kringle homology <KR4>
 F:506-583/Domain: kringle homology <KR5>
 F:620-697/Domain: kringle homology <KR6>
 F:726-803/Domain: kringle homology <KR7>
 F:840-917/Domain: kringle homology <KR8>
 F:954-1031/Domain: kringle homology <KR9>
 F:1068-1145/Domain: kringle homology <KR10>
 F:1191-1413/Domain: trypsin homology <TRY>
 Query Match 19.2%; Score 161; DB 2; Length 1420;
 Best Local Similarity 40.9%; Pred. No. 1.2e-06;
 Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;
 QY 50 CYEGNGHFYRGKASTDTMGRPCLPWNATVLTQ--QTYHAHRSDALQLGLGKKNYCRNPDN 107
 Db 1068 CYHNGQSYRGTFSTVTGRTQCQSSMTPHQKRTPENHPNDLTM-----NYCRNEDA 1122
 QY 108 RRRPWCYVQVGLKPLVQE--CMVHDCAD 133
 Db 1123 DTGPMCFCT---MDPSVRREYCNLTQCS 1147
 RESULT 27
 C61545
 plasmin (EC 3.4.21.7) precursor - goat (fragments)
 N:Alternate names: plasminogen
 C:Species: Capra aegagrus hircus (domestic goat)
 C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
 C:Accession: C61545
 R:Schaller, J.; Rickli, E.E.
 Enzyme 40, 63-69, 1988
 A:Title: Structural aspects of the plasminogen of various species.
 A:Reference number: A61545; MUID:89005015; PMID:3168975
 A:Accession: C61545
 A>Status: Preliminary
 A:Molecule type: protein
 A:Residues: 1-123 <SCH>
 C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolc
 C:Keywords: hydrolase; serine proteinase
 F:41-118/Domain: kringle homology <KR4>
 Query Match 19.1%; Score 160; DB 2; Length 123;
 Best Local Similarity 34.5%; Pred. No. 1.8e-07;
 Matches 38; Conservative 17; Mismatches 35; Indels 20; Gaps 6;
 QY 35 KKFGQC-----HCEIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNATV--LQOYTH 85
 Db 20 KKLGRSVEDCAKCE--EADQCYHNGQSYRGSTTTVGRKQCSWSSMTPHRHQKTP 78
 QY 86 AHRSDALQLGLGKKNYCRNPNRRPWCYVOVGLKPLV--QECMVHDCAD 133
 Db 79 SYPNAGLTW-----NYCRNPDAKSPWCYT---TDPRVRWEFCNLKKCE 120
 RESULT 28
 B30848
 plasmin (EC 3.4.21.7) precursor - rhesus macaque
 C:Species: Macaca mulatta (rhesus macaque)
 C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 22-Jun-1999
 C:Accession: B32869; B30848
 R:Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.
 J. Biol. Chem. 264, 5957-5965, 1989
 A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
 A:Reference number: A32869; MUID:89174660; PMID:2925643

A/Accession: B32869
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-810 <TM>
C/Cross-references: GB:J04697; NID:G342272; PIDN:AAA36901.1; PID:G342273
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:1-9/Domain: signal sequence #status predicted <SIG>
F:103-181/Domain: kringle homology <K1>
F:185-262/Domain: kringle homology <K2>
F:275-352/Domain: kringle homology <K3>
F:377-454/Domain: kringle homology <K4>
F:481-560/Domain: kringle homology <K5>
F:581-803/Domain: trypsin homology <TRY>
F:49-73, 53-61, 103-181, 124-164, 152-176, 185-262, 188-316, 206-245, 234-257, 275-352, 296-335, 32
bonds: #status predicted
F:622,665,760/Active site: His, Asp, Ser #status predicted

Query Match 19.1%; Score 159.5; DB 2; Length 810;
Best Local Similarity 31.4%; Pred. No. 1e-06;
Matches 49; Conservative 12; Mismatches 66; Indels 29; Gaps 7;

QY 10 NCCLNGGTCVSKYFNTHWNCPPKFGQ-----HCEIDKSKTCYEGNGH 56
DB 428 NPADKGPWCFTDPVRWEYCNKKCSGTGSAAPPPVAQLPDAETPSEDCMFGNGK 487

QY 57 FYRGKASTDTMGPRCLPMSATVLOQTYHAHR-----SDALQLGLGKHNYCRNPD-NRRRP 111
DB 486 GYRGKATTVTGTPCQEMAA-----QEPHSRIFTETNPRAGLEK-NYCRNPDGDVGCP 541

QY 112 WCYVQVGLKPLVQECMVHCA-----DGKXPSPPEE 143
DB 542 WCYT-TNPKRLFDYCDVPOCAASSFCGKRPQVEPKK 576

RESULT 29
E61545
Plasmin (EC 3.4.21.7) precursor - dog (fragments)
N/Alternate names: plasminogen
C/Species: Canis lupus familiaris (dog)
C/Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
C/Accession: E61545
R/Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A/Title: Structural aspects of the plasminogen of various species.
A/Reference number: A61545; MUID:8905015; PMID:3168975
A/Status: preliminary
A/Accession: E61545
A/Molecule type: protein
A/Residues: 1-120 <SCH>
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C/Keywords: hydrolase; serine proteinase
F:37-114/Domain: kringle homology <KR4>

Query Match 19.0%; Score 159; DB 2; Length 120;
Best Local Similarity 38.9%; Pred. No. 2.2e-07;
Matches 37; Conservative 8; Mismatches 32; Indels 18; Gaps 5;

QY 46 KSKTCYEGNGHYPYRGKASTDTMGPRCLPMSATVLOQTYHAHRSDAL-----QLGLGKH 100
DB 33 KYQECYHNGQSYRGTSITITGRKCSQSSMT-----PHRHKTPHFPEAGL-TMN 84

QY 101 YCRNPDNRRPWCYVQVGLKPLV--QECMVHDCAD 133
DB 85 YCRNPDADKSPCYT---TDPVSRWEFCNLRKCLD 116

RESULT 30
PUBO
Plasmin (EC 3.4.21.7) precursor - bovine
N/Alternate names: plasminogen
C/Species: Bos primigenius taurus (cattle)

A/Accession: B32869
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-810 <TM>
C/Cross-references: GB:J04697; NID:G342272; PIDN:AAA36901.1; PID:G342273
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:1-9/Domain: signal sequence #status predicted <SIG>
F:103-181/Domain: kringle homology <K1>
F:185-262/Domain: kringle homology <K2>
F:275-352/Domain: kringle homology <K3>
F:377-454/Domain: kringle homology <K4>
F:481-560/Domain: kringle homology <K5>
F:581-803/Domain: trypsin homology <TRY>
F:49-73, 53-61, 103-181, 124-164, 152-176, 185-262, 188-316, 206-245, 234-257, 275-352, 296-335, 32
bonds: #status predicted
F:622,665,760/Active site: His, Asp, Ser #status predicted

Query Match 19.1%; Score 159.5; DB 2; Length 810;
Best Local Similarity 31.4%; Pred. No. 1e-06;
Matches 49; Conservative 12; Mismatches 66; Indels 29; Gaps 7;

QY 10 NCCLNGGTCVSKYFNTHWNCPPKFGQ-----HCEIDKSKTCYEGNGH 56
DB 428 NPADKGPWCFTDPVRWEYCNKKCSGTGSAAPPPVAQLPDAETPSEDCMFGNGK 487

QY 57 FYRGKASTDTMGPRCLPMSATVLOQTYHAHR-----SDALQLGLGKHNYCRNPD-NRRRP 111
DB 486 GYRGKATTVTGTPCQEMAA-----QEPHSRIFTETNPRAGLEK-NYCRNPDGDVGCP 541

QY 112 WCYVQVGLKPLVQECMVHCA-----DGKXPSPPEE 143
DB 542 WCYT-TNPKRLFDYCDVPOCAASSFCGKRPQVEPKK 576

RESULT 29
E61545
Plasmin (EC 3.4.21.7) precursor - dog (fragments)
N/Alternate names: plasminogen
C/Species: Canis lupus familiaris (dog)
C/Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
C/Accession: E61545
R/Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A/Title: Structural aspects of the plasminogen of various species.
A/Reference number: A61545; MUID:8905015; PMID:3168975
A/Status: preliminary
A/Accession: E61545
A/Molecule type: protein
A/Residues: 1-120 <SCH>
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C/Keywords: hydrolase; serine proteinase
F:37-114/Domain: kringle homology <KR4>

Query Match 19.0%; Score 159; DB 2; Length 120;
Best Local Similarity 38.9%; Pred. No. 2.2e-07;
Matches 37; Conservative 8; Mismatches 32; Indels 18; Gaps 5;

QY 46 KSKTCYEGNGHYPYRGKASTDTMGPRCLPMSATVLOQTYHAHRSDAL-----QLGLGKH 100
DB 33 KYQECYHNGQSYRGTSITITGRKCSQSSMT-----PHRHKTPHFPEAGL-TMN 84

QY 101 YCRNPDNRRPWCYVQVGLKPLV--QECMVHDCAD 133
DB 85 YCRNPDADKSPCYT---TDPVSRWEFCNLRKCLD 116

RESULT 30
PUBO
Plasmin (EC 3.4.21.7) precursor - bovine
N/Alternate names: plasminogen
C/Species: Bos primigenius taurus (cattle)

C/Date: 30-Sep-1987 #sequence_revision 28-Apr-1995 #text_change 18-Jun-1999
C/Accession: S45046; A25835; I45961; S03736
R/Berglund, L.; Andersen, M.D.; Petersen, T.E.
submitted to the EMBL Data Library, May 1994
A/Description: Cloning and characterization of the bovine plasminogen cDNA.
A/Reference number: S45046
A/Accession: S45046
A/Molecule type: mRNA
A/Residues: 1-812 <BER>
A/Cross-references: EMBL:X79402; NID:G494962; PIDN:CAA55939.1; PID:G494963
A/Experimental source: liver
A/Note: It is uncertain whether Met-1 or Met-8 is the initiator
R/Schaller, J.; Moser, P.W.; Danneberg-Muller, G.A.K.; Rosselet, S.J.; Kampfer, U.; Ri-
bur, J. Biochem. 149, 267-278, 1985
A/Title: Complete amino acid sequence of bovine plasminogen. Comparison with human pla-
A/Reference number: A25835; MUID:85203906; PMID:3846532
A/Accession: A25835
A/Molecule type: protein
A/Residues: 27-334, 'D', 336-515, 'H', 517-554, 'L', 556-812 <SCH>
R/Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A/Title: Characterization of a complementary deoxyribonucleic acid coding for human an-
A/Reference number: I45961; MUID:85023311; PMID:6148961
A/Accession: I45961
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 706-743, 'R', 745-812 <MAL>
A/Cross-references: GB:K02935; NID:G163551; PIDN:AAA30714.1; PID:G163552
R/Brunihold, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M
Eur. J. Biochem. 114, 465-470, 1981
A/Title: Comparison of the primary structure of the N-terminal CNBr fragments of human
A/Reference number: S03735; MUID:81212097; PMID:7238497
A/Accession: S03736
A/Molecule type: protein
A/Residues: 27-83 <BRU>
C/Function:
A/Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen a
A/Pathway: fibrinolysis
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolo-
C/Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasm
F:1-26/Domain: signal sequence #status predicted <SIG>
F:8-103/Domain: plasminogen-related protein precursor homology <PLPH>
F:27-812/Product: plasminogen #status experimental <PRO>
F:27-103/Domain: activation peptide #status experimental <APT>
F:104-583, 584-812/Product: plasmin #status experimental <MAT>
F:104-583/Domain: plasmin chain A #status experimental <ACH>
F:110-188/Domain: kringle homology <K1>
F:192-269/Domain: kringle homology <K2>
F:282-359/Domain: kringle homology <K3>
F:384-461/Domain: kringle homology <K4>
F:485-564/Domain: kringle homology <K5>
F:584-812/Domain: plasmin chain B #status experimental <BCH>
F:584-805/Domain: trypsin homology <TRY>
F:56-80, 60-68, 110-188, 131-171, 159-183, 192-269, 195-323, 213-252, 241-264, 282-359, 303-342,
bonds: #status predicted
F:315/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:365/Binding site: carbohydrate (Ser) (covalent) #status experimental
F:624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 19.0%; Score 159; DB 1; Length 812;
Best Local Similarity 28.0%; Pred. No. 1.1e-06;
Matches 49; Conservative 15; Mismatches 55; Indels 56; Gaps 11;

QY 5 HQVPSNCDCLN-----GGTCVSKYFNTHW--CNCPKFGQGHCHIDSKT 49
DB 315 NRTPEFPCKNLEENYCRNPNKXAPWCYTTN--SEVRWEYCTIPS-----CSSPLST 366

QY 50 -----CYEGNGHYRGKASTDTMGPRCLPMSATVLOQTYHAH----RS 89
DB 367 ERMDVPVPPEQTPTVQDCYHNGQSGTGTSTITGRKCSQSSWSS-----MTPHRLKTP 421

QY 90 DALQLGLGKHNYCRNPDNRRPWCYVQVGLKPLV--QECMVHDCADGK--PSSP 140

Db 422 NYPNAGL-TWNYCRNPADKSPWCYT---TDPVRWFEFCNLKKCSETPQVPAAP 472

Search completed: December 3, 2003, 14:44:18
Job time : 7.39655 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003. Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:53 ; Search time 4.59483 Seconds
(without alignments)
1463.563 Million cell updates/sec

Title: US-09-880-503-8

Perfect score: 837

Sequence: 1 SNELHQVPSNCDLNGTGV.....QECMVHDCADGKPKSPPEE 143

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	837	100.0	431	1	UROK_HUMAN
2	757.5	90.5	433	1	P00749 homo sapien
3	625.5	74.7	442	1	P16227 papio cynoc
4	505	72.3	433	1	P04185 sus scrofa
5	592	70.7	432	1	Q05589 bos taurus
6	569	68.0	433	1	P29598 rattus norv
7	335.5	40.1	477	1	P06869 mus musculu
8	334.5	40.0	434	1	P15638 desmodus ro
9	328.5	39.2	431	1	P15120 gallus gall
10	328.5	39.2	562	1	P98121 desmodus ro
11	325	38.8	559	1	P00750 homo sapien
12	315.5	37.7	559	1	P19637 rattus norv
13	310.5	37.1	477	1	P11214 mus musculu
14	300.5	35.9	566	1	P98119 desmodus ro
15	268	32.0	603	1	Q28198 bos taurus
16	262	31.3	655	1	Q04962 cavia porce
17	250.5	29.9	653	1	Q04756 homo sapien
18	233	27.8	615	1	Q9-098 mus musculu
19	223	26.6	394	1	P00748 homo sapien
20	217	25.9	593	1	P49150 desmodus ro
21	169	20.2	685	1	P98140 bos taurus
22	162.5	19.4	4548	1	Q24488 drosophila
23	161	19.2	473	1	P08519 homo sapien
24	161	19.2	1420	1	Q92484 rattus norv
25	159.5	19.1	810	1	P14117 macaca mula
26	159	19.0	452	1	P12545 macaca mula
27	159	19.0	473	1	Q90Y90 xenopus lae
28	159	19.0	812	1	P06868 bos taurus
29	158.5	18.9	937	1	Q01973 homo sapien
30	158.5	18.9	937	1	Q92139 mus musculu
31	156.5	18.7	810	1	P00747 homo sapien
32	156	18.6	475	1	Q96m8 homo sapien
33	155	18.5	790	1	P06867 sus scrofa

ALIGNMENTS

RESULT 1

ID	UROK_HUMAN	STANDARD	PRT;	431 AA.
AC	P00749; Q15844; Q15618; Q969M6;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	20-MAR-1987 (Rel. 04, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)			
DE	(U-plasminogen activator)			
GN	PLAU.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85215647; PubMed=2987867;			
RA	Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Biasi F.;			
RT	"The human urokinase-plasminogen activator gene and its promoter.";			
RL	Nucleic Acids Res. 13:2759-2771(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,			
RT	Steffens G.J., Heyneker H.L.;			
RT	"Cloning and expression of the gene for pro-urokinase in Escherichia coli.";			
RL	Biotechnology 3:923-929(1985).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86056954; PubMed=2415429;			
RA	Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,			
RA	Nishida M., Suyama T.;			
RT	"Molecular cloning of cDNA coding for human prepro-urokinase.";			
RL	Gene 36:183-188(1985).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85203359; PubMed=3888571;			
RA	Jacobs P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P.,			
RA	van Elsen A., Herzog A., Bollen A.;			
RT	"Molecular cloning, sequencing, and expression in Escherichia coli of human prepro-urokinase cDNA.";			
RL	DNA 4:139-146(1985).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RA	Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,			
RA	Nickerson D.A.;			
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RX	TISSUE=Lung;			
RA	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Atschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore H., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			

34	155	18.5	812	1	PLMN_MOUSE	P20918 mus musculu
35	153.5	18.3	810	1	PLMN_ERIEU	Q29485 erinaceus e
36	152.5	18.2	169	1	PLMN_RAT	Q01177 rattus norv
37	150	17.9	711	1	HGFL_HUMAN	P26927 homo sapien
38	148.5	17.7	462	1	KRM2_HUMAN	Q8ncw0 homo sapien
39	148	17.7	333	1	PLMN_CANFA	P80009 canis famil
40	148	17.7	943	1	ROR2_HUMAN	Q01974 homo sapien
41	148	17.7	944	1	ROR2_MOUSE	Q92118 mus musculu
42	147.5	17.6	716	1	HGFL_MOUSE	P26928 mus musculu
43	146.5	17.5	461	1	KRM2_MOUSE	Q8k1s7 mus musculu
44	144.5	17.3	728	1	HGF_HUMAN	P14210 homo sapien
45	142	17.0	728	1	HGF_MOUSE	Q08048 mus musculu

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahay J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7].
 RP SEQUENCE OF 66-431 FROM N.A.
 RX MEDLINE=84272706; PubMed=6589620;
 RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.,
 RT "Identification and primary sequence of an unspliced human urokinase
 RT poly(A)+ RNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
 RN [8].
 RP SEQUENCE OF 21-177.
 RX MEDLINE=83055084; PubMed=6754569;
 RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
 RT Flohe L.;
 RT "The primary structure of high molecular mass urokinase from human
 RT urine. The complete amino acid sequence of the A chain.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
 RN [9].
 RP SEQUENCE OF 156-176 AND 179-224.
 RX MEDLINE=83003608; PubMed=6749491;
 RA Schaller J., Nick H., Rickli E.E., Gillesen D., Lergier W.,
 RA Studer R.O.;
 RT "Human low-molecular-weight urinary urokinase. Partial
 RT characterization and preliminary sequence data of the two polypeptide
 RT chains.";
 RL Eur. J. Biochem. 125:251-257(1982).
 RN [10].
 RP SEQUENCE OF 158-410.
 RX MEDLINE=83055099; PubMed=6754572;
 RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.;
 RT "The complete amino acid sequence of low molecular mass urokinase
 RT from human urine.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
 RN [11].
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=96000858; PubMed=8591045;
 RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
 RA Dobson C.M., Stuart D.I., Jones E.Y.;
 RT "The crystal structure of the catalytic domain of human
 RT urokinase-type plasminogen activator.";
 RL Structure 3:681-691(1995).
 RN [12].
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
 RX MEDLINE=20266327; PubMed=10805774;
 RA Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
 RA Bode W., Magdolen V., Huber R., Moroder L.;
 RT "[4-aminomethyl]phenylguanidine derivatives as nonpeptidic highly
 RT selective inhibitors of human urokinase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
 RN [13].
 RP STRUCTURE BY NMR.
 RX MEDLINE=89127526; PubMed=2536903;
 RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;
 RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
 RT dimensional NMR.";
 RL Nature 337:579-582(1989).
 RN [14].
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=93003110; PubMed=1327118;
 RA Li X., Smith R.A.G., Dobson C.M.;

"Sequential 1H NMR assignments and secondary structure of the kringle
 domain from urokinase.";
 RL Biochemistry 31:9562-9571(1992).
 RN [15].
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=94149701; PubMed=8107091;
 RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;
 RT "Solution structure of the kringle domain from urokinase-type
 RT plasminogen activator.";
 RL J. Mol. Biol. 235:1548-1559(1994).
 RN [16].
 RP VARIANT LEU-141.
 RX MEDLINE=96186279; PubMed=8652631;
 RA Yoshimoto M., Ushiyama Y., Sakai M., Tanaki S., Hara H., Takahashi K.,
 RA Sawasaki Y., Hanada K.;
 RT "Characterization of single chain urokinase-type plasminogen
 RT activator with a novel amino-acid substitution in the kringle
 RT structure.";
 RL Biochim. Biophys. Acta 1293:83-89(1996).
 RN [17].
 RP VARIANT LEU-141.
 RX MEDLINE=97218551; PubMed=9065988;
 RA Conne B., Berczy M., Belin D.;
 RT "Detection of polymorphisms in the human urokinase-type plasminogen
 RT activator gene.";
 RL Thromb. Haemost. 77:434-435(1997).
 RN [18].
 RP ERRATUM.
 RA Conne B., Berczy M., Belin D.;
 RL Thromb. Haemost. 78:973-973(1997).
 RN [19].
 RP VARIANT LEU-141.
 RX MEDLINE=973737920; PubMed=9194591;
 RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W.,
 RA Creutzburg S., Graeff H., Magdolen V.;
 RT "Mutation analysis of the genes encoding urokinase-type plasminogen
 RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
 RL Electrophoresis 18:686-689(1997).
 CC -!- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR
 CC THERAPY OF THROMBOLYTIC DISORDERS.
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
 CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
 CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
 CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.
 CC -!- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used
 CC in Pulmonary Embolism (PE) to initiate fibrinolysis.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; X02419; CAA26268.1; -
 DR EMBL; M15476; AAA61253.1; -
 DR EMBL; D00244; BAA00175.1; -
 DR EMBL; D11143; BAA01919.1; -
 DR EMBL; X02760; CAA26535.1; -
 DR EMBL; AF377330; AAK53822.1; -
 DR EMBL; BC013575; AAH13575.1; -
 DR EMBL; K03226; AAC97138.1; -
 DR EMBL; K02286; AAA61252.1; -
 DR EMBL; A21571; CAA01559.1; -
 DR EMBL; A18397; CAA01390.1; -
 DR PIR; A00931; UKGHU.
 DR PDB; 1KDU; 31-OCT-93.

Query Match 100.0%; Score 837; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1.2e-73;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCCLNGTGVSNKYFNIHWCNPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
 DB 21 SNELHQPNSCCLNGTGVSNKYFNIHWCNPKFKGGQHCIEDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGWRPCLPWNATVLOQTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120
 DB 81 KASTDTMGWRPCLPWNATVLOQTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGKPPSPPEE 143
 DB 141 PLVQECMVHDCADGKPPSPPEE 163

RESULT 2
 UROK_PAPCY STANDARD; PRT; 433 AA.

AC P16227;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.

OS Papio cynocephalus (Yellow baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheinae; Papio.
 OX NCBI_TaxID=9556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thoracic aorta;
 RX MEDLINE=90287734; PubMed=2113276;
 RA AU Y.P.T., Wang T.W., Clowes A.W.;
 RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen activator";
 RL Nucleic Acids Res. 18:3411-3411(1990).
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.
 CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 CC EMBL; X51935; CAA36200.1; -;
 CC FIR; S14687; UKBAY.
 CC HSSP; P00749; LLMW.
 CC MEROPS; S01.231; -;
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR006209; EGF like.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00051; kringle; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000395; Kringle; 1.

DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE NEG.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein; Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 176 CHAIN A (BY SIMILARITY)
 FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
 FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
 FT DOMAIN 26 62 EGF-LIKE.
 FT DOMAIN 69 150 KRINGLE.
 FT DOMAIN 151 177 CONNECTING PEPTIDE.
 FT DOMAIN 178 433 SERINE PROTEASE.
 FT DISULFID 30 38 BY SIMILARITY.
 FT DISULFID 32 50 BY SIMILARITY.
 FT DISULFID 52 61 BY SIMILARITY.
 FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 208 224 BY SIMILARITY.
 FT DISULFID 216 287 BY SIMILARITY.
 FT DISULFID 315 384 BY SIMILARITY.
 FT DISULFID 347 363 BY SIMILARITY.
 FT DISULFID 374 402 BY SIMILARITY.
 FT ACT_SITE 223 223 CHARGE RELAY SYSTEM.
 FT ACT_SITE 274 274 CHARGE RELAY SYSTEM.
 FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
 FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).
 SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

Query Match 90.5%; Score 757.5; DB 1; Length 433;
 Best Local Similarity 92.3%; Pred. No. 5.7e-66;
 Matches 132; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 SNELHQPNSCCLNGTGVSNKYFNIHWCNPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
 DB 21 SREL-QVPSDCGLNGTGVSNKYFNIHWCNPKFKGGQHCIEDKSKTCYEGNGHFYRG 79

QY 61 KASTDTMGWRPCLPWNATVLOQTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120
 DB 80 KASTDTMGWRPCLPWNATVLOQTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 139

QY 121 PLVQECMVHDCADGKPPSPPEE 143
 DB 140 QRVQECMVHDCADGKPPSPPEE 162

RESULT 3
 UROK_PIG STANDARD; PRT; 442 AA.

AC P04185;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU.

OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9923;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=85087954; PubMed=6096832;
 RX Negamine Y., Pearson D., Altus M.S., Reich E.;
 RA "cDNA and gene nucleotide sequence of porcine plasminogen activator";
 RT

RL Nucleic Acids Res. 12:9525-9541 (1984).
RN [2].
RP REVISION TO 241.
RA Nagamine Y.;
RL Submitted (DEC-1986) to the PIR data bank.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01648; CAA25806.1; -;
DR EMBL; X02724; CAA26511.1; -;
DR PIR; A00932; UKPG.
DR HSP; P00749; IKDU.
DR MEROPS; S01.231; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE-NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00020; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 442
FT CHAIN 21 188
FT CHAIN 190 442
FT DOMAIN 29 65
FT DOMAIN 72 153
FT DOMAIN 154 189
FT DOMAIN 190 442
FT CARBOHYD 152 152
FT DISULFID 33 41
FT DISULFID 35 53
FT DISULFID 55 64
FT DISULFID 179 310
FT DISULFID 220 236
FT DISULFID 228 299
FT DISULFID 324 393
FT DISULFID 356 372
FT DISULFID 383 411
FT ACT_SITE 235 235
FT ACT_SITE 286 286
FT ACT_SITE 387 387
FT ACT_SITE 241 241
FT CONFLICT 242 242
FT CONFLICT 286 286
FT CONFLICT 288 288
FT SEQUENCE 442 AA; 49116 MW; E332FCBF501321EE CRC64;
Query Match 74.7%; Score 625.5; DB 1; Length 442;
Best Local Similarity 73.4%; Pred. No. 3.2e-53;
Matches 113; Conservative 11; Mismatches 19; Indels 11; Gaps 2;

QY 1 SHELQV--PSNCDCLNGGTQVSNKYFNIHMCNCPKFGQGHCEIDSKSKTCYENGHFY 58
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
21 SHELQESGASNCGLNGKCVSYKFSNIQRCSCPKEFGHECEIDTSQTCFEGNGHSY 80
QY 59 RGAASDTMTWGRPCLPWSATVLQOYTHAHRSDALQGLGKKNYCNPNRNPWCYVQVG 118
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
81 RGAANTGTGRPCLPWSATVLLNTYHAHRPDALQGLGKKNYCNPNRNPWCYVQVG 140
QY 119 LKPLVQECMVDHCA-----DGKFPSPPEE 143
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
141 LKQLVQECMVPNCSGESHRPAYDGKNPFSTPEK 174
RESULT 4
UROK_BOVIN STANDARD; PRT; 433 AA.
ID UROK BOVIN Q05589; Q28209;
AC Q05589; Q28209;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aortic endothelium;
RX MEDLINE=93216119; PubMed=8385052;
RA Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,
RA Schleuning W.-D.;
RT "Bovine urokinase-type plasminogen activator and its receptor:
RT cloning and induction by retinoic acid.";
RL Gene 125:177-183(1993).
RN [2]
RP SEQUENCE OF 12-433 FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RT and tPA.";
RL Int. Dairy J. 5:605-617(1995).
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- INDUCTION: By retinoic acid.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L03546; AAA51419.1; -;
DR EMBL; X85801; CAA59796.1; -;
DR PIR; JN0560; JN0560.
DR HSP; P00749; ILMW.
DR MEROPS; S01.231; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.

```

DR DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00020; TRYD_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00022; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 433
FT CHAIN 21 179
FT CHAIN 181 433
FT DOMAIN 29 65
FT DOMAIN 72 133
FT DOMAIN 154 180
FT DOMAIN 181 433
FT DISULFID 33 41
FT DISULFID 35 53
FT DISULFID 55 64
FT DISULFID 170 301
FT DISULFID 211 227
FT DISULFID 219 290
FT DISULFID 315 384
FT DISULFID 347 363
FT DISULFID 374 402
FT ACT_SITE 226 226
FT ACT_SITE 277 277
FT ACT_SITE 378 378
FT CONFLICT 189 189
FT SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;

Query Match 72.3%; Score 605; DB 1; Length 433;
Best Local Similarity 72.4%; Pred. No. 3e-51;
Matches 105; Conservative 14; Mismatches 24; Indels 2; Gaps 1;

QY 1 SHELQV--PSNCDLNGTCTVSNKYFNSHNWCNPKFGQGHCEIDSKTCYENGHGF 58
DB |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
21 SNEVHKESESNCGLNGKGVCTYKTYFNSIQRCSCPKFKFGQGHCEIDSKTCYQNGHGF 80
QY 59 RKGASTDTMGRCPLPNSATVLTQOTVHAHRSALQLGLGKHNYCNPDRRPPWCYQVVG 118
DB |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
81 RKGANDLSGRCLAWDSPTVLKHYHAHRSALQLGLGKHNYCNPDRRPPWCYQVVG 140
QY 119 LKPLVQECMVHDCADGKPSPEE 143
DB |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
141 LKQFVQFCMVQDCSVGKSPSPREK 165

RESULT 5
UROK_RAT STANDARD; PRT; 432 AA.
AC P29598;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-SEP-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer 344;
RX MEDLINE=92233409; PubMed=1568219;
RA Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,
RA Kefferd R.F.;
RT "Transcriptional and posttranscriptional activation of urokinase

```

```

RT Plasminogen activator gene expression in metastatic tumor cells.;
Cancer Res. 52:2489-2496(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Rabbani S.A.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC Plasminogen to form Plasmin.
CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH
CC CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM
CC CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE
CC LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63434; CAA45028.1; -.
DR EMBL; X65651; CAA46601.1; -.
DR PIR; S24604; S18932.
DR HSSP; P00749; 1KDU.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 19
FT CHAIN 20 432
FT CHAIN 20 177
FT CHAIN 156 177
FT CHAIN 179 432
FT DOMAIN 27 63
FT DOMAIN 70 151
FT DOMAIN 152 178
FT DOMAIN 179 432
FT DISULFID 31 39
FT DISULFID 33 51
FT DISULFID 53 62
FT DISULFID 168 300
FT DISULFID 210 226
FT DISULFID 218 289
FT DISULFID 314 383
FT DISULFID 346 362
FT DISULFID 373 401
FT ACT_SITE 225 225
FT ACT_SITE 276 276
FT ACT_SITE 377 377

```

```

MGI:97611; Plau.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01386; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS03400; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Plasmogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 178 CHAIN A (BY SIMILARITY).
FT CHAIN 157 178 SHORT A CHAIN (A1).
FT CHAIN 180 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 28 64 EGF-LIKE.
FT DOMAIN 71 152 KRINGLE.
FT DOMAIN 153 179 CONNECTING PEPTIDE.
FT DOMAIN 180 433 SERINE PROTEASE.
FT DISULFID 32 40 BY SIMILARITY.
FT DISULFID 34 52 BY SIMILARITY.
FT DISULFID 54 63 BY SIMILARITY.
FT DISULFID 169 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM.
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
T SEQUENCE 433 AA; 48268 MW; A99C35F6250443F9 CRC64;
Query Watch 68.0%; Score 569; DB 1; Length 433;
Best Local Similarity 70.4%; Pred. No. 9.le-48;
Matches 95; Conservative 15; Mismatches 25; Indels 0; Gaps 0
QY 9 SNCDCLNGGTCVSNKYFSNIHWCNPKKFGQGHCEIDKSKTCYEGNGHFGYRGKASTDTMG 68
Db 30 SNGCGQGGVCVSVKYFSIRRCSCPRKFGHEHCEIDASKTCYHGNGDSYRGKANTDTG 89
QY 69 RCLPWNATVLQTYHAHRSDALQLGLGHXNYCRNPDNRRRPMWCYVQGLKPLVOECMV 128
Db 90 RCLAWNAPVLQPKYNAHRPDAISLGLGHXNYCRNPDNQRKPMWCYVQGLRQFVOECMV 149
QY 129 HDCADGKKPSPPEE 143
Db 150 HDCSLSKKPSSSVDQ 164
RESULT 7
URT2_DESRO
ID URT2_DESRO STANDARD; PRT; 477 AA.
AC P15638;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSFA
DE alpha-2) (BAR-PA) (I-plasminogen activator).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; Desmodus.
 OX NCBI_TaxID=9430;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=92039036; PubMed=1937019;
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RA Alagon A., Donner P., Schleuning W.D.;
 RA "The plasminogen activator family from the salivary gland of the
 RT vampire bat Desmodus rotundus: cloning and expression.";
 RL Gene 105:229-237(1991).
 RN [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Salivary gland;
 RX MEDLINE=9003667; PubMed=2509450;
 RA Cardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
 RA Regester R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;
 RA "Isolation, characterization, and cDNA cloning of a vampire bat
 RT salivary plasminogen activator.";
 RL J. Biol. Chem. 264:17947-17952(1989).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=93393059; PubMed=1309059;
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kraetzschmar J., Haendler B., Langer G., Baidus B., Witt W.,
 RA Donner P.;
 RA "plasminogen activators from the saliva of Desmodus rotundus (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
 CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
 CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
 CC AGENT.
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
 CC THE PRESENCE OF FIBRIN I.
 CC -!- SUBUNIT: Monomer.
 CC -!- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
 CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
 CC STIMULATION OF ACTIVITY.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 1 kringle domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; M63988; AAA31593.1; --
 CC EMBL; J05082; AAA31596.1; --
 CC PIR; J05098; J05098.
 CC HSP; P98119; IAS1.
 CC MEROPS; S01.232; --
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR006209; EGF like.
 CC InterPro; IPR000083; Fibrinctn.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00008; EGF; 1.
 CC Pfam; PF00039; fn1; 1.
 CC Pfam; PF00051; kringle; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000395; Kringle; 1.

DR SMART; SMO0181; EGF; 1.
 DR SMART; SMO0058; FN1; 1.
 DR SMART; SMO0130; KR; 1.
 DR SMART; SMO0020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Signal; Multigene family.
 FT SIGNAL 1 36 POTENTIAL.
 FT CHAIN 37 477 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 2.
 FT DOMAIN 40 82 FIBRONECTIN TYPE-I.
 FT DOMAIN 83 121 EGF-LIKE.
 FT DOMAIN 128 209 KRINGLE.
 FT DOMAIN 225 477 SERINE PROTEASE.
 FT ACT_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 321 321 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 428 428 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 42 72 BY SIMILARITY.
 FT DISULFID 70 79 BY SIMILARITY.
 FT DISULFID 87 98 BY SIMILARITY.
 FT DISULFID 92 109 BY SIMILARITY.
 FT DISULFID 111 120 BY SIMILARITY.
 FT DISULFID 128 209 BY SIMILARITY.
 FT DISULFID 149 191 BY SIMILARITY.
 FT DISULFID 180 204 BY SIMILARITY.
 FT DISULFID 214 345 BY SIMILARITY.
 FT DISULFID 257 273 BY SIMILARITY.
 FT DISULFID 265 334 BY SIMILARITY.
 FT DISULFID 359 434 BY SIMILARITY.
 FT DISULFID 391 407 BY SIMILARITY.
 FT DISULFID 424 452 BY SIMILARITY.
 FT CARBOHYD 185 185 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 403 403 N -> K (IN REF. 2).
 FT CONFLICT 417 417 Y -> H (IN REF. 2).
 FT CONFLICT 435 435 M -> R (IN REF. 2).
 SQ SEQUENCE 477 AA; 53719 MW; 17486555C0B5077C CRC64;
 Query Match 40.1%; Score 335.5; DB 1; Length 477;
 Best Local Similarity 46.3%; Pred. No. 3.5e-25;
 Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;
 Qy 3 ELHQVP-----SNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCEIDKSKTCYEGNGHFY 58
 Db 78 QCHTVPVKSCSELRCFNGGTCWQAASFSDF-VCCPKGYTGKQCEVDTHATCYKDGVTY 136
 Qy 59 RGASTDTRGRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRPNDRRRRRCWYQVQG 118
 Db 137 RGWTSSESQAQINNNNLLTRTYNGRRSDAITLGLGNHNYCRPNDRNNKRPWCYVKA 196
 Qy 119 LKPLVQECWVHDC A 132
 Db 197 SKTILFCSVPVCS 210

RESULT 8

UROK_CHICK
 ID UROK_CHICK STANDARD; PRT; 434 AA.
 AC P5120;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.
RN NCBI_TaxID=9031;
RP SEQUENCE FROM N.A.
RX MEDLINE=90110185; PubMed=2295632;
RA Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;
RT "The chicken urokinase-type plasminogen activator gene.";
RL J. Biol. Chem. 265:1339-1344(1990).
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; J05187; AAA49131.1; -;
DR ENBL; J05188; AAA49130.1; -;
DR PIR; A35005; A35005.
DR HSP; P00763; IDPO.
DR MEROPS; S01.231; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01196; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 434
FT CHAIN 21 171
FT CHAIN 173 434
FT DOMAIN 36 72
FT DOMAIN 79 158
FT DOMAIN 159 172
FT DOMAIN 173 434
FT DOMAIN 40 60
FT DISULFID 42 60
FT DISULFID 62 71
FT DISULFID 162 296
FT DISULFID 202 218
FT DISULFID 210 285
FT DISULFID 310 379
FT DISULFID 342 358
FT DISULFID 365 397
FT ACT_SITE 217 217
FT ACT_SITE 272 272
FT ACT_SITE 373 373
FT CARBOHYD 228 228
SQ SEQUENCE 434 AA; 49400 MW; BD881048DD666A55 CRC64;
Query Match 40.0%; Score 334.5; DB 1; Length 434;

Best Local Similarity 54.2%; Pred. No. 4e-25; Mismatches 14; Indels 5; Gaps 2;
Matches 58; Conservative 14; Mismatches 30; Indels 5; Gaps 2;
QY 11 CDCLINGGTCVSNKYFSNIHWCNCPKPGQHCIDKSKTCYEGNGHFYRGKASDTMGRP 70
DB 40 COCLNGGTCIITYRFSQIKRCLCPGEGYGLHCEIDTNSICYSGNGEDYRGMAEDP---G 95
QY 71 CLPWNASIVLQ-QYVHAHRSALQGLGKHNKYNCRPNDRRPPWCYVQ 116
DB 96 CLYNDHPSVIRMGDYHADLKNALQGLGKHNKYNCRPNDRRPPWCYTK 142
RESULT 9
URTB DESRO
ID URTB DESRO STANDARD; PRT; 431 AA.
AC P98121;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA
DE beta).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; M63989; AAA31594.1; -;
DR PIR; J50599; J50599.
DR HSP; P98119; IAS1.
DR MEROPS; S01.239; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00051; kringle; 1.

```

DR Pfam: PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00016; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp SPC; 1.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS01186; EGF 2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 431
FT DOMAIN 37 75
FT DOMAIN 82 163
FT DOMAIN 179 431
FT ACT_SITE 226 226
FT ACT_SITE 275 275
FT ACT_SITE 382 382
FT DISULFID 41 52
FT DISULFID 46 63
FT DISULFID 65 74
FT DISULFID 82 163
FT DISULFID 103 145
FT DISULFID 134 158
FT DISULFID 168 299
FT DISULFID 211 227
FT DISULFID 219 288
FT DISULFID 313 388
FT DISULFID 345 361
FT DISULFID 378 406
FT CARBOHYD 139 139
FT CARBOHYD 352 352
FT SEQUENCE 431 AA; 48221 MW; 69985675B162CBF CRC64;

Query Match 39.2%; Score 328.5; DB 1; Length 431;
Best Local Similarity 47.6%; Pred. No. 1.5e-24;
Matches 59; Conservative 16; Mismatches 48; Indels 1; Gaps 1;

QY 9 SNCDCLNGTCVSNKYFNSIHWCNCPKFGQHCEIDKSKTYEGNGHYPRGKASTDTWG 68
DB 42 SELRCFNGGTCWQAASFDF-VQCCKPGYTGKQCEVDTHATCYDQGVYRGWTSTSSG 100
QY 69 RPLCPWNSATVLTQTYAHRSDALQLGLGKHNYCRNPDRRBPWCYVQVGLKPLVQECMV 128
DB 101 AQCIWNSNLLTRTYNGRSDAILTGLGNHNYCRNPDRNNSPWCYVVKASKFIIEFCSV 160
QY 129 HDCA 132
DB 161 PVCS 164

RESULT 10
TPA_HUMAN
ID TPA_HUMAN STANDARD; PRT; 562 AA.
AC P00750; Q15103;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator) (Alteplase) (Retelase).
GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE-Melanoma;
RX MEDLINE=83115262; PubMed=6337343;
RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA Goeddel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RL cDNA in E. coli.";
RL Nature 301:214-221(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=88262579; PubMed=3133640;
RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT from human fetal lung cells.";
RL Nucleic Acids Res. 16:5695-5695(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,
RA Hsiung N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RT mouse cells using BPV vectors.";
RL DNA 6:461-472(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196143; PubMed=3009482;
RA Frieznar Degen S.J., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene.";
RL J. Biol. Chem. 261:6972-6985(1986).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=84298137; PubMed=6089198;
RA Ny T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RT correlation of intron and exon structures to functional and
RT structural domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Entage J.S.,
RA Oopenaker G., Volckaert G., Rombauts W., Billiau A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RT and its expression in Escherichia coli.";
RL Mol. Biol. Med. 3:279-292(1986).
RN [7]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Umbilical vein;
RX MEDLINE=90192129; PubMed=2107528;
RA Siebert P.D., Fong K.;
RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
RT human endothelial cells.";
RL Nucleic Acids Res. 18:1086-1086(1990).
RN [8]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Colling F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

```


SEQUENCE FROM N.A.
 MEDLINE=89170114; PubMed=3148445;
 NY T., Leonardson G., Haueh A.J.W.;
 "Cloning and characterization of a cDNA for rat tissue-type
 plasminogen activator.";
 DNA 7:671-677(1988).
 (2)
 SEQUENCE FROM N.A.
 MEDLINE=90130448; PubMed=2105315;
 Feng P., Ohlsson M., Ny T.;
 "The structure of the TATA-less rat tissue-type plasminogen activator
 gene. Species-specific sequence divergences in the promoter predict
 differences in regulation of gene expression.";
 J. Biol. Chem. 265:2022-2027(1990).
 CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
 TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 MANY OTHER PHYSIOLOGICAL EVENTS.
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 plasminogen to form plasmin.
 CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
 BOND.
 CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
 CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
 PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
 ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
 CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
 CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 2 kringle domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 EMBL; M23697; AAA4182.1; -;
 EMBL; M31197; AAA42261.1; -;
 EMBL; M31185; AAA42261.1; JOINED.
 EMBL; M31186; AAA42261.1; JOINED.
 EMBL; M31187; AAA42261.1; JOINED.
 EMBL; M31188; AAA42261.1; JOINED.
 EMBL; M31189; AAA42261.1; JOINED.
 EMBL; M31190; AAA42261.1; JOINED.
 EMBL; M31191; AAA42261.1; JOINED.
 EMBL; M31192; AAA42261.1; JOINED.
 EMBL; M31193; AAA42261.1; JOINED.
 EMBL; M31194; AAA42261.1; JOINED.
 EMBL; M31195; AAA42261.1; JOINED.
 EMBL; M31196; AAA42261.1; JOINED.
 EMBL; A19618; CAA01482.1; -;
 PIR; A35029; A35029.
 HSP; P00750; 1RTP.
 MEROPS; S01.232; -;
 InterPro; IPR001314; Chymotrypsin.
 InterPro; IPR006209; EGF-like.
 InterPro; IPR000083; Fibronectin.
 InterPro; IPR006210; IEGF.
 InterPro; IPR000001; Kringle.
 InterPro; IPR001254; Ser_protease_Try.
 Pfam; PF00008; EGF; 1.
 Pfam; PF00039; fn1; 1.
 Pfam; PF00051; kringle; 2.
 Pfam; PF00089; trypsin; 1.
 PRINTS; PR00722; CHYMOTRYPSIN.
 PRINTS; PR00018; KRINGLE.
 ProDom; PD000395; Kringle; 2.

DR SMART; SM00181; EGF; 1.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00021; KRINGLE_1; 2.
 DR PROSITE; PS00070; KRINGLE_2; 2.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
 FT SIGNAL 1 17
 FT PROPEP 18 29
 FT CHAIN 30 559
 FT CHAIN 30 308
 FT CHAIN 309 559
 FT CHAIN 36 78
 FT DOMAIN 79 117
 FT DOMAIN 124 205
 FT DOMAIN 213 294
 FT DOMAIN 309 559
 FT ACT_SITE 355 355
 FT ACT_SITE 404 404
 FT ACT_SITE 510 510
 FT DISULFID 38 68
 FT DISULFID 66 75
 FT DISULFID 83 94
 FT DISULFID 88 105
 FT DISULFID 107 116
 FT DISULFID 124 205
 FT DISULFID 145 187
 FT DISULFID 176 200
 FT DISULFID 213 294
 FT DISULFID 234 276
 FT DISULFID 265 289
 FT DISULFID 297 428
 FT DISULFID 340 356
 FT DISULFID 348 417
 FT DISULFID 442 516
 FT DISULFID 474 490
 FT DISULFID 506 534
 FT CARBOHYD 149 149
 FT CARBOHYD 481 481
 FT CONFLICT 380 380
 SQ SEQUENCE 559 AA; 62903 MW; 7DBD3809C1D1C921 CRC64;
 Query Match 38.8%; Score 325; DB 1; Length 559;
 Best Local Similarity 44.1%; Pred. No. 4.2e-24;
 Matches 64; Conservative 15; Mismatches 56; Indels 10; Gaps 3;
 Qy 3 ELHQVP-----SNCDCLNGGTCVSNKYFSNIHMCNCPKFGGHCEDKSKTCYEGNGHFY 58
 Db 74 QCHSVPRVSCSBPCFNGGTCQALFSDF-VCCPDGFGKGCIDTRATCEGGQITY 132
 Qy 59 RGKASTDTMGRPCLPMNSATVLCQTYHAHRSDALQLGLGHNYCRNPNRRRRPWCYQVG 118
 Db 133 RGTWSTAENGAEICINWNSSALSQKPYARRPNAIKLGLGHNYCRNPNDRDVKPCYVKA 192
 Qy 119 LKPLVQECMVHDCADGKYPSSPPEE 143
 Db 193 GKYTFECSTPAC-----PKGPTD 212
 RESULT 12
 ID TPA_MOUSE
 AC P11214; Q91VP2; PRT; 559 AA.
 DT 01-JUL-1989 (Rel. 11, Created)

DR	MEROPS, S01.232; .-
DR	MGB; MGI:97610; Plat.
DR	InterPro; IPR001314; Chymotrypsin.
DR	InterPro; IPR006209; EGF-like.
DR	InterPro; IPR000083; Fibronectn.
DR	InterPro; IPR006210; tEGF.
DR	InterPro; IPR000001; Kringle.
DR	InterPro; IPR001254; Ser_protease_Try.
DR	Pfam; PF00008; EGF; 1.
DR	Pfam; PF00039; fnl; 1.
DR	Pfam; PF00051; kringle; 2.
DR	Pfam; PF00089; trypsin; 1.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	PRINTS; PR00018; KRINGLE.
DR	ProDom; PD000395; kringle; 2.
DR	SMART; SM00181; EGF; 1.
DR	SMART; SM00058; FN1; 1.
DR	SMART; SM00130; KR; 2.
DR	SMART; SM00020; Tryp_spc; 1.
DR	D.A. PROSITE; PS00022; EGF_1; 1.
DR	PROSITE; PS01186; EGF_2; 1.
DR	PROSITE; PS01253; FIBRONECTIN_1; 1.
DR	PROSITE; PS00021; KRINGLE_1; 2.
DR	PROSITE; PS50070; KRINGLE_2; 2.
DR	PROSITE; PS50240; TRYPSIN_DOM; 1.
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.
DR	PROSITE; PS00135; TRYPSIN_SER; 1.
KW	Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW	plasma; Kringle; EGF-like domain; Repeat; Signal.
KW	SIGNAL
FT	1 17 PROBABLE.
FT	PROPEP 18 29
FT	CHAIN 30 559
FT	CHAIN 30 308
FT	TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT	CHAIN.
FT	TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT	CHAIN.
FT	FIBRONECTIN TYPE-I.
FT	EGF-LIKE.
FT	KRINGLE 1.
FT	KRINGLE 2.
FT	SERINE PROTEASE.
FT	CHARGE RELAY SYSTEM.
FT	CHARGE RELAY SYSTEM.
FT	CHARGE RELAY SYSTEM.
FT	ACT SITE 510 510
FT	DISULFID 38 68
FT	DISULFID 66 75
FT	DISULFID 83 94
FT	DISULFID 88 105
FT	DISULFID 107 116
FT	DISULFID 124 205
FT	DISULFID 145 187
FT	DISULFID 176 200
FT	DISULFID 213 294
FT	DISULFID 234 276
FT	DISULFID 265 289
FT	DISULFID 297 428
FT	DISULFID 340 356
FT	DISULFID 348 417
FT	DISULFID 442 516
FT	DISULFID 474 490
FT	DISULFID 506 534
FT	CARBOHYD 149 149
FT	CARBOHYD 481 481
FT	CONFLICT 260 260
FT	CONFLICT 325 325
FT	SEQUENCE 559 AA; 63122 MW; 8CCSE2BDB94514D9 CRC64;
SR	Query Match 37.7%; Score 315.5; DB 1; Length 559;
SR	Best Local Similarity 44.5%; Pred.No. 3.5e-23;
SR	Matches 61; Conservative 15; Mismatches 56; Indels 5;
SR	3 ELHOVP----SNCDCLNGCTGVNKNYFSNIHWNCNPKKFGGQHCEIDYSKTCYCGEN

Db 74 QCHSVFVRSCEPRFCNGTCCQALYFSDP-VCQCPDGVGKRCDDTRATCFEQQITY 132
QY 59 RGKASTDTWGRCLPWNSTLVLOQYHARSALOLGLGKHNCRPNRPPWCYQVG 118
Db 133 RGTWTAESGACINWNSVLSKLPYNARPNALKLGHNHNYCRNPRDRLKPCYVFK 192
QY 119 LKPLVQECVMWHDCAQK 135
Db 193 GKYTFEFCSTPACPKG 209

RESULT 13

URTL DESRO
ID URTL DESRO STANDARD; PRT; 477 AA.
AC P98119;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSFA
alpha-1).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]_TaxID=9430;
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RC TISSUE=Salivary gland;
RX MEDLINE=98022741; PubMed=9354616;
RA Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,
RA Schleuning W.D., Bode W.;
RT "Catalytic domain structure of vampire bat plasminogen activator: a
RT molecular paradigm for proteolysis without activation cleavage.";
RL Biochemistry 36:13483-13493(1997).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. POTENT THROMBOLYTIC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
CC THE PRESENCE OF FIBRIN I.
CC -!- SUBUNIT: Monomer.
CC -!- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
CC STIMULATION OF ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.
CC EMBL; M63987; AAA31591.1; -
CC EMBL; M63986; AAA31592.1; -
CC PIR; JS0597; JS0597.
CC PDB; 1A5I; 23-MAR-99.
CC MEROPS; S01.232; -
CC GlycoSuiteDB; P98119; -
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR000083; Fibrinctnl.
CC InterPro; IPR006210; tEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00033; fn1; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00058; FN1; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01253; FIBRONECTIN_1; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00075; KRINGLE_2; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
CC Kringle; EGF-like domain; Signal; Multigene family; 3D-structure.
FT SIGNAL 1 36
FT CHAIN 37 477
FT DOMAIN 40 82
FT DOMAIN 83 121
FT DOMAIN 128 209
FT DOMAIN 225 477
FT ACT_SITE 272 272
FT ACT_SITE 321 321
FT ACT_SITE 428 428
FT DISULFID 42 72
FT DISULFID 70 79
FT DISULFID 87 98
FT DISULFID 92 109
FT DISULFID 111 120
FT DISULFID 128 209
FT DISULFID 149 191
FT DISULFID 180 204
FT DISULFID 214 345
FT DISULFID 257 273
FT DISULFID 265 334
FT DISULFID 359 434
FT DISULFID 391 407
FT DISULFID 424 452
FT CARBOHYD 153 153
FT CARBOHYD 398 398
FT TURN 214 215
FT STRAND 223 224
FT TURN 226 227
FT STRAND 230 231
FT HELIX 234 236
FT TURN 238 239
FT STRAND 240 245
FT STRAND 254 263
FT TURN 264 265
FT STRAND 266 269
FT HELIX 271 273
N-LINKED (GLCNAC...)
/FTid-CAR_000027.
N-LINKED (GLCNAC...)
/FTid-CAR_000028.

```

FT TURN 280 282
FT STRAND 284 287
FT TURN 297 298
FT STRAND 300 309
FT TURN 311 312
FT TURN 315 317
FT TURN 319 320
FT STRAND 323 328
FT STRAND 338 338
FT TURN 339 340
FT STRAND 341 341
FT STRAND 345 345
FT TURN 349 350
FT TURN 355 356
FT STRAND 358 363
FT STRAND 366 366
FT STRAND 374 374
FT STRAND 379 385
FT HELIX 388 390
FT TURN 393 398
FT TURN 403 404
FT STRAND 405 409
FT TURN 425 426
FT TURN 428 429
FT STRAND 431 436
FT TURN 437 438
FT STRAND 439 448
FT TURN 455 456
FT STRAND 459 463
FT HELIX 464 467
FT HELIX 468 474
SQ SEQUENCE 477 AA; 53616 MW; AA06FD1739C10E5E CRC64;

Query Match 37.1%; Score 310.5; DB 1; Length 477;
Best Local Similarity 45.5%; Pred. No. 9.1e-23;
Matches 60; Conservative 14; Mismatches 53; Indels 5; Gaps 3;

QY 5 HOVPSN-CD---CLNGGTCVSNKYFSNIHWCNCPKFGGHOCEIDKSKTCYEGNGHYRG 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 80 HTVPNSCSBPRFNGTCHQAVYFSDP-VCCPAGYTGRCVEDTRATCYEGGVYTRG 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 KASTDTWGRCLPWNSTVILQOYTHAHRSDALQGLGKHNYCRPNRRAPWCYVQVGLK 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 139 TWSTAERVCINWSSLLTRTYNGRMPDAFNGLGNHNYCRNPNGAPKPCWCYVIKAGK 198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 121 PLVQECMVHDC A 132
DB 199 FTSESCSVFVCS 210

RESULT 14
TPA_BOVIN STANDARD; PRT; 566 AA.
AC Q28198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue-Kidney;
RA Rayn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators upa
RL Int. Dairy J. 5:605-617(1995).
CC -I- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN

```

```

CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -I- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -I- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -I- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -I- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -I- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -I- SIMILARITY: Contains 1 EGF-like domain.
CC -I- SIMILARITY: Contains 1 fibronectin type I domain.
CC -I- SIMILARITY: Contains 2 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X85800; CAA59795.1; -.
CC HSP; P00750; IRTF.
CC MEROPS; S01.232.
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR006209; EGF-like.
CC InterPro: IPR000083; FibrinctnI.
CC InterPro: IPR006210; IEFG.
CC InterPro: IPR000001; Kringle.
CC InterPro: IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00039; fnl; 1.
CC Pfam; PF00051; kringle; 2.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 2.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00058; FNI; 1.
CC SMART; SM00130; KR; 2.
CC SMART; SM00020; Tryp_SPC; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01253; FIBRONECTIN_1; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 2.
CC PROSITE; PS0240; TRYPsin_DOM; 1.
CC PROSITE; PS00134; TRYPsin_HIS; 1.
CC PROSITE; PS00135; TRYPsin_SER; 1.
CC PLASMA; Kringle; EGF-like domain; Repeat; Signal.
KW plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW SIGNAL.
FT PROPEP 1 21 BY SIMILARITY.
FT CHAIN 22 33 BY SIMILARITY.
FT CHAIN 34 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 34 314 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 315 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT DOMAIN 40 82 FIBRONECTIN TYPE-I.
FT DOMAIN 83 121 EGF-LIKE.
FT DOMAIN 128 209 KRINGLE 1.
FT DOMAIN 219 300 KRINGLE 2.
FT DOMAIN 315 566 SERINE PROTEASE.
FT ACT_SITE 361 361 CHARGE RELAY SYSTEM.
FT ACT_SITE 410 410 CHARGE RELAY SYSTEM.
FT ACT_SITE 517 517 CHARGE RELAY SYSTEM.
FT DISULFID 42 72 BY SIMILARITY.

```

FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 87 98 BY SIMILARITY.
FT DISULFID 92 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 128 209 BY SIMILARITY.
FT DISULFID 149 191 BY SIMILARITY.
FT DISULFID 180 204 BY SIMILARITY.
FT DISULFID 219 300 BY SIMILARITY.
FT DISULFID 240 282 BY SIMILARITY.
FT DISULFID 271 295 BY SIMILARITY.
FT DISULFID 303 434 INTERCHAIN (BY SIMILARITY).
FT DISULFID 346 362 BY SIMILARITY.
FT DISULFID 354 423 BY SIMILARITY.
FT DISULFID 448 523 BY SIMILARITY.
FT DISULFID 480 496 BY SIMILARITY.
FT DISULFID 513 541 BY SIMILARITY.
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 566 AA; 63701 MW; 2EB6BEB4E32276C3 CRC64;

Query Match 35.9%; Score 300.5; DB 1; Length 566;
Best Local Similarity 44.0%; Pred. No. 9.9e-22;
Matches 59; Conservative 16; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQP-SNCD---CLNGTGVSNKYFNIHWCNCPKFGGQHCEIDKSKTCYEGNGHPT 58
DB 78 QCHSVFVRSCEPWCFCNGTGRQALYSDF-VQCPEGFMGKLCIEDATATCYKDGQVAY 136
QY 59 RGKASTDTWGRPCLPNWATVLQCTVHAHRSALQLGLGKHVCRNPDRRPPWCYQVVG 118
DB 137 RGTWSTAESGAECANWSSGLAMKPGSRPNAILGLGNHVCNPDODSRPWCYVFKA 196
QY 119 LKPLVQECMWHQCA 132
DB 197 GKYISEFCSTPACA 210

RESULT 15
ID FA12_CAVPO STANDARD; PRT; 603 AA.
AC Q04962;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAUF) (Fragment).
GN F12.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.
RC TISSUE=Liver;
RX MEDLINE=93003367; PubMed=1390917;
RA Semba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,
RA Kambara T., Okabe H.;
RT Primary structure of guinea-pig Hageman factor: sequence around the
RT cleavage site differs from the human molecule."
RL Biochim. Biophys. Acta 1159:113-121 (1992).
CC -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-ile bonds in factor
CC VII to form factor VIIa and factor XI to form factor Xla.
CC -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.

CC CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC CC -!- SIMILARITY: Contains 1 kringle domain.
CC CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X68615; CAA48600.1; -
DR PIR; S28941; S28941.
DR HSP; P00763; 1DPO.
DR MEROPS; S01_211; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; K-ingle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
FT NON_TER 1 1
FT SIGNAL <1 18
FT CHAIN 19 358 ALPHA-FACTOR XIIA HEAVY CHAIN.
FT CHAIN 359 603 ALPHA-FACTOR XIIA LIGHT CHAIN.
FT DOMAIN 46 87 FIBRONECTIN TYPE-II.
FT DOMAIN 93 130 EGF-LIKE 1.
FT DOMAIN 132 172 FIBRONECTIN TYPE-I.
FT DOMAIN 173 209 EGF-LIKE 2.
FT DOMAIN 216 294 KRINGLE.
FT DOMAIN 312 342 PRO-RICH.
FT DOMAIN 359 603 SERINE PROTEASE.
FT ACT_SITE 398 398 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 551 551 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 97 109 BY SIMILARITY.
FT DISULFID 103 118 BY SIMILARITY.
FT DISULFID 120 129 BY SIMILARITY.
FT DISULFID 134 162 BY SIMILARITY.
FT DISULFID 160 169 BY SIMILARITY.
FT DISULFID 177 188 BY SIMILARITY.
FT DISULFID 182 197 BY SIMILARITY.
FT DISULFID 199 208 BY SIMILARITY.
FT DISULFID 216 294 BY SIMILARITY.

```

FT DISULFID 237 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 345 472 BY SIMILARITY.
FT DISULFID 383 399 BY SIMILARITY.
FT DISULFID 391 461 BY SIMILARITY.
FT DISULFID 422 425 BY SIMILARITY.
FT DISULFID 488 557 BY SIMILARITY.
FT DISULFID 520 536 BY SIMILARITY.
FT DISULFID 547 578 BY SIMILARITY.
FT CARBOHYD 248 248 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLNAC. . .) (POTENTIAL).
SQ SEQUENCE 603 AA; 66795 MW; 48DC6B946FB9ED59 CRC64;

Query Match 32.0%; Score 268; DB 1; Length 603;
Best Local Similarity 38.6%; Pred. No. 1.5e-18;
Matches 51; Conservative 21; Mismatches 52; Indels 8; Gaps 3;

QY 13 CLNGGTCVSNKYNIHWCNPKKFGQHCIDKSKTCYEGNHGFRGKASTDTMGRCPL 72
Dbb 182 CLNGGRCLE---VEGHLCDPCMGYTGPFCDLDTTASCYEGRGVSYRGMMARTTVSGAKCQ 238
QY 73 PWSATVLQOTYHAHRSD-ALQGLGKHNYCRPNDRRREWCYVQGLKELVQECNVHDC 131
Dbb 239 RWAS----EATYRNMTAEQALRGHLGHTFCRPNDRTPWCFVWGNRLSWEYCDLAQC 294
QY 132 ADGKKSPPEE 143
Dbb 295 QYPPQPTATPHD 306

RESULT 16
HGFA HUMAN
ID HGFA_HUMAN STANDARD; PRT; 655 AA.
AC Q04756; Q14726;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver, and Serum;
RX MEDLINE=93252878; PubMed=7683665;
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
RA Kitamura N.
RT "Molecular cloning and sequence analysis of the cDNA for a human
RT serine protease responsible for activation of hepatocyte growth
RT factor. Structural similarity of the protease precursor to blood
RT coagulation factor XII."
RL J. Biol. Chem. 268:10024-10028(1993).
RN [2]
RP SEQUENCE OF 40-655 FROM N.A.
RA Zhao S., Odell C.
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY
CC CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
CC -1- TISSUE SPECIFICITY: LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D14012; BAA03113.1; -
DR EMBL; Z69923; CAA93803.1; -
DR PIR; A46688; A46688.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.228; -
DR Gensw; HGNC:4894; HGFA.
DR MIM; 604552; -
DR GO; GO:0005576; C:extracellular; TAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR007042; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF-2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 30
FT PROPEP 31 372 CLEAVED IN ACTIVE FORM.
FT CHAIN 373 407 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT
FT CHAIN 408 655 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG
FT CHAIN 108 148 FIBRONECTIN TYPE-II.
FT DOMAIN 160 198 EGF-LIKE 1.
FT DOMAIN 200 240 FIBRONECTIN TYPE-I.
FT DOMAIN 241 279 EGF-LIKE 2.
FT DOMAIN 286 367 KRINGLE.
FT DOMAIN 408 655 SERINE PROTEASE.
FT ACT_SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 497 497 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 598 598 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 108 133 BY SIMILARITY.
FT DISULFID 122 148 BY SIMILARITY.
FT DISULFID 164 175 BY SIMILARITY.
FT DISULFID 169 186 BY SIMILARITY.
FT DISULFID 188 197 BY SIMILARITY.
FT DISULFID 202 230 BY SIMILARITY.

```

FT DISULFID 228 237 BY SIMILARITY.
FT DISULFID 245 256 BY SIMILARITY.
FT DISULFID 250 267 BY SIMILARITY.
FT DISULFID 269 278 BY SIMILARITY.
FT DISULFID 286 367 BY SIMILARITY.
FT DISULFID 307 349 BY SIMILARITY.
FT DISULFID 338 362 BY SIMILARITY.
FT DISULFID 394 521 INTERCHAIN (BY SIMILARITY).
FT DISULFID 432 448 BY SIMILARITY.
FT DISULFID 440 510 BY SIMILARITY.
FT DISULFID 535 604 BY SIMILARITY.
FT DISULFID 567 583 BY SIMILARITY.
FT DISULFID 594 622 BY SIMILARITY.
FT CARBOHYD 48 48 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 492 492 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 546 546 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 644 644 R -> Q (IN REF. 2).
SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1E1B862ED7 CRC64;

Query Match 31.3%; Score 262; DB 1; Length 655;
Best Local Similarity 36.9%; Pred. No. 6e-18;
Matches 58; Conservative 12; Mismatches 57; Indels 30; Gaps 3;

QY 5 HQVPSNCDCLNGTGVSNKYFSNIHW-----CNCPKKFGGQHCIEDKSKTCYEGNGH 56
DB 242 HTACLSPPCLNGTGC-----HLIVATGTTVCACPGFAGRLCNTEPDERCFLNGT 292
QY 57 FYRGKASTDTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVYQ 116
DB 293 GYRGVASTSAGLSCLAWNSDLYQLHVDVSGAALGLGPHAYCRNPDNDERPCVYV 352
QY 117 VGLKPLVQECMVHDC-----ADGKKPSP 140
DB 353 KDSALSWEYCRLEACESLTVQLSPDLLATLPEPASP 389

RESULT 17
HGFA_MOUSE
ID HGFA_MOUSE STANDARD; PRT; 653 AA.
AC Q9R098; Q9JUV4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF activator) (HGFA).
GN HGFA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Itoh H., Kataoka H., Koono H.;
RT "Mouse hepatocyte growth factor activator."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J., Yang J., Huan Y.;
RT "Activation of HGF by endogenous HGF activator is required for metanephric kidney morphogenesis in vitro."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
EMBL; AF090917; AAF02489.1; -
EMBL; AF224724; AAF34712.1; -
HSPF; P00763; LDPO.
MEROPS; S01.228; -
MGD; MGI:1859281; Hgfac.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01185; EGF_2; 1.
DR PROSITE; PS01253; FIBONECTIN_1; 1.
DR PROSITE; PS00023; FIBONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29 BY SIMILARITY.
FT PROPEP 30 369 CLEAVED IN ACTIVE FORM (BY SIMILARITY).
FT CHAIN 370 405 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT CHAIN.
FT CHAIN 406 653 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG CHAIN.
FT DOMAIN 105 145 FIBONECTIN TYPE-II.
FT DOMAIN 157 195 EGF-LIKE 1.
FT DOMAIN 197 237 FIBONECTIN TYPE-I.
FT DOMAIN 238 276 EGF-LIKE 2.
FT DOMAIN 283 364 KRINGLE.
FT DOMAIN 406 653 SERINE PROTEASE.
FT ACT_SITE 445 445 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 495 495 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 596 596 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 105 130 BY SIMILARITY.
FT DISULFID 119 145 BY SIMILARITY.
FT DISULFID 161 172 BY SIMILARITY.
FT DISULFID 166 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 199 227 BY SIMILARITY.
FT DISULFID 225 234 BY SIMILARITY.

CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
 CC FACTOR XII TO FORM KALLIKREIN WHICH THEN CLEAVES FACTOR XII FIRST
 CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
 CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -|- SIMILARITY: Contains 2 EGF-like domains.
 CC -|- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -|- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -|- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M31315; AAA70225.1; -;
 CC DR EMBL; AF538691; AAM97932.1; -;
 CC DR EMBL; M11723; AAA51986.1; -;
 CC DR EMBL; M17466; AAB59490.1; -;
 CC DR EMBL; M17464; AAB59490.1; JOINED.
 CC DR EMBL; M17465; AAB59490.1; JOINED.
 CC DR EMBL; M13147; AAA70224.1; -;
 CC DR EMBL; U71274; AAB51203.1; -;
 CC DR PIR; A29411; KFHU12.
 CC DR HSSP; P00763; IDPO.
 CC DR MEROPS; S01.211; -;
 CC DR Gensu; HGNC:3530; F12.
 CC DR MIM; 234000; -;
 CC GO; GO:0003805; F:blood coagulation factor XI activity; TAS.
 CC DR GO; GO:0003806; F:blood coagulation factor XII activity; TAS.
 CC DR GO; GO:0007596; F:blood coagulation; TAS.
 CC DR InterPro; IPR001314; Chymotrypsin.
 CC DR InterPro; IPR006209; EGF like.
 CC DR InterPro; IPR000083; Fibrinctnl.
 CC DR InterPro; IPR000562; FN_Type_II.
 CC DR InterPro; IPR006210; IEGF.
 CC DR InterPro; IPR000001; Kringle.
 CC DR InterPro; IPR001254; Ser_protease_Try.
 CC DR Pfam; PF000008; EGF; 2.
 CC DR Pfam; PF000039; fn1; 1.
 CC DR Pfam; PF000040; fn2; 1.
 CC DR Pfam; PF000051; kringle; 1.
 CC DR Pfam; PF000089; trypsin; 1.
 CC DR PRINTS; PR00722; CHYMOTRYPSIN.
 CC DR PRINTS; PR00013; FNTYPEII.
 CC DR PRINTS; PR00018; KRINGLE.
 CC DR ProDom; PD000995; FN_Type_II; 1.
 CC DR ProDom; PD000195; Kringle; 1.
 CC DR SMART; SM00181; EGF; 2.
 CC DR SMART; SM000059; FN1; 1.
 CC DR SMART; SM000059; FN2; 1.
 CC DR SMART; SM00130; KR; 1.
 CC DR SMART; SM00020; Tryp_SPC; 1.
 CC DR PROSITE; PS00022; EGF_1; 2.
 CC DR PROSITE; PS01186; EGF_2; 1.
 CC DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 CC DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 CC DR PROSITE; PS00021; KRINGLE_1; 1.
 CC DR PROSITE; PS00070; KRINGLE_2; 1.
 CC DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC DR PROSITE; PS00135; TRYPSIN_SER; 1.
 CC KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
 CC KW Hydrolase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;
 CC KW Polymorphism; Disease mutation.
 CC FT SIGNAL 1 19
 CC FT CHAIN 20 372 ALPHA-FACTOR XIIA HEAVY CHAIN.
 CC FT CHAIN 373 615 ALPHA-FACTOR XIIA LIGHT CHAIN.
 CC FT CHAIN 354 362 BETA-FACTOR XIIA PART 1.
 CC FT CHAIN 373 615 BETA-FACTOR XIIA PART 2.

FT DOMAIN 47 88 FIBRONECTIN TYPE-II.
 FT DOMAIN 94 131 EGF-LIKE 1.
 FT DOMAIN 133 173 FIBRONECTIN TYPE-I.
 FT DOMAIN 174 210 EGF-LIKE 2.
 FT DOMAIN 217 295 KRINGLE.
 FT DOMAIN 296 349 PRO-RICH.
 FT DOMAIN 373 615 SERINE PROTEASE.
 FT CARBOHYD 109 109 O-LINKED (FUC).
 FT CARBOHYD 249 249 N-LINKED (GLCNAC...).
 FT CARBOHYD 299 299 O-LINKED (POTENTIAL).
 FT CARBOHYD 305 305 O-LINKED (POTENTIAL).
 FT CARBOHYD 308 308 O-LINKED (POTENTIAL).
 FT CARBOHYD 328 328 O-LINKED (POTENTIAL).
 FT CARBOHYD 329 329 O-LINKED (POTENTIAL).
 FT CARBOHYD 337 337 O-LINKED (POTENTIAL).
 FT ACT_SITE 412 412 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 461 461 CHARGE RELAY SYSTEM (BY SIMILARITY).
 Query Match 27.8%; Score 233; DB 1; Length 615;
 Best Local Similarity 42.3%; Pred. No. 3.6e-15;
 Matches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;
 Qy 13 CLNGGTCVSNKYFSNIHWCNPKKFGQHCIEIDKSTCYEGNGHFGKASTDTMRPCL 72
 Db 183 CLHGRCLE---VEGHRLLCHCPVGYTGFCDVTKASCYDGRGLSYRGLARTTLSPGAPCQ 239
 Qy 73 PWSATVLQOQY-HAHRSDALQLGLGKHNCRPNDRRRPCVY 115
 Db 240 PWAS----EATYRNVTAEQARNWGLGHAFCRAPDNDIRPWCVF 279
 RESULT 19
 URUG DESRO
 ID URUG DESRO STANDARD; PRT; 394 AA.
 AC P49150;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 28-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSFA gamma).
 DE OS Desmodus rotundus (Vampire bat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
 OC Desmodontinae; Desmodus.
 OC NCBI_TaxID=9430;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland.
 RX MEDLINE=92039036; PubMed=1937019;
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
 RA Alagon A., Donner P., Schleuning W.D.;
 RT "The plasminogen activator family from the salivary gland of the
 RT vampire bat Desmodus rotundus: cloning and expression.";
 RL Gene 105:229-237(1991).
 RL [2]
 RP CHARACTERIZATION.
 RX MEDLINE=9393059; PubMed=1309059;
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
 RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
 RA Donner P.;
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common
 RT vampire bat): unique fibrin specificity.";
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
 CC -|- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
 CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
 CC AGENT.
 CC -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -|- SUBUNIT: Monomer.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -|- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration

DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
 KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal;
 FT NON_TER 1 1
 FT SIGNAL <1 9 POTENTIAL.
 FT CHAIN 10 349 ALPHA-FACTOR X1IA HEAVY CHAIN.
 FT CHAIN 350 593 ALPHA-FACTOR X1IA LIGHT CHAIN.
 FT DOMAIN 37 78 FIBRONECTIN TYPE-II.
 FT DOMAIN 84 121 EGF-LIKE 1.
 FT DOMAIN 123 163 FIBRONECTIN TYPE-I.
 FT DOMAIN 164 200 EGF-LIKE 2.
 FT DOMAIN 207 287 KRINGLE.
 FT DOMAIN 297 333 PRO-RICH.
 FT DOMAIN 350 593 SERINE PROTEASE.
 FT ACT_SITE 389 399 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 438 438 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 541 541 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 88 100 BY SIMILARITY.
 FT DISULFID 94 109 BY SIMILARITY.
 FT DISULFID 111 120 BY SIMILARITY.
 FT DISULFID 125 153 BY SIMILARITY.
 FT DISULFID 151 160 BY SIMILARITY.
 FT DISULFID 168 179 BY SIMILARITY.
 FT DISULFID 173 198 BY SIMILARITY.
 FT DISULFID 190 199 BY SIMILARITY.
 FT DISULFID 207 287 BY SIMILARITY.
 FT DISULFID 230 269 BY SIMILARITY.
 FT DISULFID 258 282 BY SIMILARITY.
 FT DISULFID 336 463 BY SIMILARITY.
 FT DISULFID 374 390 BY SIMILARITY.
 FT DISULFID 382 452 BY SIMILARITY.
 FT DISULFID 413 416 BY SIMILARITY.
 FT DISULFID 479 547 BY SIMILARITY.
 FT DISULFID 510 526 BY SIMILARITY.
 FT DISULFID 537 568 BY SIMILARITY.
 FT CARBOHYD 99 99 O-LINKED (FUC) (BY SIMILARITY).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 410 410 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 593 AA; 65148 MW; 7215928A792BD61F CRC64;
 Query Match 25.9%; Score 217; DB 1; Length 593;
 Best Local Similarity 35.7%; Pred. No. 1.2e-13;
 Matches 46; Conservative 15; Mismatches 58; Indels 10; Gaps 4;
 QY 6 QVPSNCDLNGTGVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYE--GNCHFYRGKAS 63
 Db 166 QVCRTPNPLNGSCLOAE---GRLRCRCAPSAGRLCDVLDKASCYDDRDRLGSLYRGWAG 222
 QY 64 TDTMRGRCPLPWSATVLOQTY-HAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLKPL 122
 Db 223 TILSGAPCQSWAS-----EATYNVTAQVNLWGLGDHAFCRPNPDTRPWCIFWIKGDRLS 278
 QY 123 VQECMVHDC 131
 Db 279 WNYCRLAPC 287
 RESULT 21
 ID ROR1_DROME STANDARD; PRT; 685 AA.
 AC Q24488;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase transmembrane receptor Ror precursor
 DE (EC 2.7.1.112) (ROR).
 GN ROR OR CG4926.

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
 RC STRAIN=Canton-S; TISSUE=Larval brain;
 RX MEDLINE=9334822; PubMed=8394009;
 RA Wilson C., Goberdhan D.C.I., Steller H.;
 RT "Dror, a potential neurotrophic receptor gene, encodes a Drosophila
 RT homolog of the vertebrate Ror family of Trk-related receptor tyrosine
 RL kinases";
 RN [2]
 RP Proc. Natl. Acad. Sci. U.S.A. 90:7109-7113(1993).
 RC SEQUENCE FROM N.A.
 RX STRAIN=Berkely;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Chen B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Balaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Betos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Furi V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Testor C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RN [3]
 RP Science 287:2185-2195(2000).
 RC SEQUENCE OF 545-597 FROM N.A.
 RX MEDLINE=98401146; PubMed=9731193;
 RA Oates A.C., Wollberg P., Achen M.G., Wilks A.F.;
 RT "Sampling the genomic pool of protein tyrosine kinase genes using the
 RT polymerase chain reaction with genomic DNA";
 RL Biochem. Biophys. Res. Commun. 249:660-667(1998).
 CC -!- FUNCTION: Tyrosine-protein kinase receptor that functions during
 CC early stages of neuronal development.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: Expressed in neurons of the developing nervous
 CC system.
 CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ROR
 CC SUBFAMILY.
 CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.

CC -!- SIMILARITY: Contains 1 kringle domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L20297; AAA28860.1; ..
 DR EMBL; AE003628; AAF52885.1; ..
 DR EMBL; AJ002908; CAA05743.1; ..
 DR PIR; A48289; A48289.
 DR HSP; 011362; 1FGK.
 DR FlyBase; FBgn0010407; Ror.
 DR GO; GO:0016021; C: integral to membrane; NAS.
 DR GO; GO:0004713; F: protein tyrosine kinase activity; NAS.
 DR GO; GO:0007417; F: central nervous system development; IEP.
 DR GO; GO:0006469; F: protein amino acid phosphorylation; NAS.
 DR InterPro; IPR000024; Fz domain.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002011; RukinaseII.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PRO0018; KRINGE.
 DR PRINTS; PRO0109; TYRKINASE.
 DR ProDom; PD000395; Kringle; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00130; K3; 1.
 DR SMART; SM00219; TyrK2; 1.
 DR PROSITE; PS00038; FZ; 1.
 DR PROSITE; PS00021; KRINGE_1; 1.
 DR PROSITE; PS00070; KRINGE_2; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
 DR Transferrase; Kinase; Tyrosine-protein kinase; ATP-binding; Receptor;
 KW Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
 KW Developmental protein.
 FT SIGNAL 1 24
 FT CHAIN 25 685
 FT DOMAIN 25 317
 FT TRANSMEM 318 338
 FT DOMAIN 339 585
 FT DOMAIN 36 225
 FT DOMAIN 236 310
 FT DOMAIN 410 677
 FT NP_BIND 416 424
 FT BINDING 442 442
 FT ACT_SITE 539 539
 FT MOD_RES 565 565
 FT MOD_RES 569 569
 FT MOD_RES 570 570
 FT MOD_RES 570 570
 FT CARBOHYD 45 45
 FT CARBOHYD 63 63
 FT CARBOHYD 129 129
 FT CARBOHYD 144 144
 FT CARBOHYD 250 250
 FT SEQUENCE 685 AA; 78442 MW; 526162D27D5FD7C7 CRC64;
 Query Match 20.2%; Score 169; DB 1; Length 685;
 Best Local Similarity 28.8%; Pred. No. 6e-09;
 Matches 38; Conservative 27; Mismatches 35; Indels 32; Gaps 6;
 3 ELHQVPSNCDLNGTCTVSNKYFNIHWCNCFKFGGCHCEIDKSKTCYEGNGHVRGA 62

DB 212 DCQKLPQHKCCLSGITV-----EVDKTCNYWEDGSTTRGVA 249
 QY 63 STDWGRPCLPWNSATVLTQYTHAHRSDALQLGKHNCRPN-RRRPRWCYVQGLKP 121
 DB 250 NVSASGRKPLRW--SWLMKEI-----SDPPEL-IGQ-NYCRNFGSVENSPWCFVDSRR 300
 QY 122 LVQSCMVHDCAD 133
 DB 301 ILELDIPKCAD 312
 RESULT 22
 APOA HUMAN STANDARD; PRT; 4548 AA.
 AC P08519;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Apolipoprotein(a) precursor (EC 3.4.21.-) (Apo(a)) (Lp(a)).
 GN LPA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=88039109; PubMed=3670400;
 RA McLean J.W., Tomlison J.E., Kuang W.-J., Eaton D.L., Chen E.Y.,
 RA Fless G.M., Scanu A.M., Lawn R.M.;
 RT "cDNA sequence of human apolipoprotein(a) is homologous to
 RT plasminogen.";
 RL Nature 330:132-137(1987).
 RN [2]
 RP SERINE PROTEASE ACTIVITY.
 RP MEDLINE=90076123; PubMed=2531657;
 RA Salonen E.-M., Jauhainen M., Zardi L., Vaheri A., Ehnholm C.;
 RT "Lipoprotein(a) binds to fibronectin and has serine proteinase
 RT activity capable of cleaving it.";
 RL EMBO J. 8:4035-4040(1989).
 RN [3]
 RP REVIEW.
 RP MEDLINE=90049223; PubMed=2530631;
 RA Utermann G.;
 RT "The mysteries of lipoprotein(a).";
 RL Science 246:904-910(1989).
 RN [4]
 RP CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.
 RP MEDLINE=21303595; PubMed=11294842;
 RA Garner B., Merry A.H., Royle L., Harvey D.J., Rudd P.M., Thillet J.;
 RT "Structural elucidation of the N- and O-glycans of human
 RT apolipoprotein(a): role of o-glycans in conferring protease
 RT resistance.";
 RL J. Biol. Chem. 276:22200-22208(2001).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.
 RX MEDLINE=96217891; PubMed=8642595;
 RA Mikol V., Lograsso P.V., Boettcher B.R.;
 RT "Crystal structures of apolipoprotein(a) kringle IV37 free and
 RT complexed with 6-aminohexanoic acid and with p-aminomethylbenzoic
 RT acid: existence of novel and expected binding modes.";
 RL J. Mol. Biol. 256:751-761(1996).
 RN [6]
 RP VARIANT ARG-4193.
 RP MEDLINE=95002201; PubMed=7918682;
 RA Scanu A.M., Pfaffinger D., Lee J.C., Hinman J.;
 RT "A single point mutation (Trp72-->Arg) in human apo(a) kringle 4-37
 RT associated with a lysine binding defect in Lp(a).";
 CC Biochim. Biophys. Acta 1227:41-45(1994).
 CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
 CC (Lp(a)). It has serine proteinase activity and is able of
 CC autolysis. Inhibits tissue-type plasminogen activator 1.
 CC Lp(a) may be a ligand for megalin/Gp 330.

CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
 CC decorin.
 CC -!- PM: N- and O-glycosylated. The N-glycans are complex biantennary
 CC structures present in either a mono- or disialylated state. The
 CC O-glycans are mostly (80%) represented by the monosialylated core
 CC type I structure, NeuNAcGal2-3Galbeta1-3GalNAc, with smaller
 CC amounts of disialylated and non-sialylated O-glycans also
 CC detected.
 CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its
 CC naturally occurring proteolytic fragments are correlated with
 CC atherosclerosis. Homology with plasminogen kringle IV and V is
 CC thought to underlie the atherogenicity of the protein, because the
 CC fragments are competing with plasminogen for fibrin(ogen) binding.
 CC -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
 CC leading to the formation of the so called mini-p(a). Apo(a)
 CC fragments accumulate in atherosclerotic lesions, where they may
 CC promote thrombogenesis. O-glycosylation may limit the extent of
 CC proteolytic fragmentation.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -!- SIMILARITY: Contains 38 kringle domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X06290; CAA29618.1; --
 CC PIR: S00657; S00657.
 CC PDB: 1J71; 13-JUN-01.
 CC PDB: 1JFN; 28-JUN-01.
 CC PDB: 1KIV; 18-MAY-99.
 CC PDB: 3KIV; 18-MAY-99.
 CC PDB: 4KIV; 18-MAY-99.
 CC MEROPS: S01.226; --
 CC Genew: HGNC:6667; LPA.
 CC MIM: 152200; --
 CC DR GO: GO:0004866; F: endopeptidase inhibitor activity; TAS.
 CC DR GO: GO:0008015; P: circulation; TAS.
 CC DR GO: GO:0009405; P: pathogenesis; TAS.
 CC DR InterPro: IPR001314; Chymotrypsin.
 CC DR InterPro: IPR000001; Kringle.
 CC DR InterPro: IPR001254; Ser-protease_Try.
 CC DR Pfam: PF00051; Kringle; 38.
 CC DR PRINTS: PR00722; CHYMOTRYPSIN.
 CC DR PRINTS: PR00018; KRINGLE.
 CC DR ProDom: PD000395; Kringle; 38.
 CC DR SMART: SM00130; KR; 38.
 CC DR SMART: SM00020; Tryp_SPC; 1.
 CC DR PROSITE: PS00021; KRINGLE_1; 38.
 CC DR PROSITE: PS00070; KRINGLE_2; 38.
 CC DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 CC DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 CC DR PROSITE: PS00135; TRYPSIN_SER; 1.
 CC DR Hydrolase: Serine protease; Lipid transport; Plasma; Glycoprotein;
 CC Kringle; Repeat; Atherosclerosis; Signal; Polymorphism; 3D-structure.
 CC SIGNAL 1 19
 CC CHAIN 20 4548 APOLIPOPROTEIN(A).
 CC DOMAIN 20 130 KRINGLE TYPE IV, 1.
 CC DOMAIN 131 244 KRINGLE TYPE IV, 2.
 CC DOMAIN 245 358 KRINGLE TYPE IV, 3.
 CC DOMAIN 359 472 KRINGLE TYPE IV, 4.
 CC DOMAIN 473 586 KRINGLE TYPE IV, 5.
 CC DOMAIN 587 700 KRINGLE TYPE IV, 6.
 CC DOMAIN 701 814 KRINGLE TYPE IV, 7.
 CC DOMAIN 815 928 KRINGLE TYPE IV, 8.
 CC DOMAIN 929 1042 KRINGLE TYPE IV, 9.
 CC DOMAIN 1043 1156 KRINGLE TYPE IV, 10.
 CC DOMAIN 1157 1270 KRINGLE TYPE IV, 11.
 CC DOMAIN 1271 1384 KRINGLE TYPE IV, 12.

FT DOMAIN 1385 1498 KRINGLE TYPE IV, 13.
 FT DOMAIN 1499 1612 KRINGLE TYPE IV, 14.
 FT DOMAIN 1613 1726 KRINGLE TYPE IV, 15.
 FT DOMAIN 1727 1840 KRINGLE TYPE IV, 16.
 FT DOMAIN 1841 1954 KRINGLE TYPE IV, 17.
 FT DOMAIN 1955 2068 KRINGLE TYPE IV, 18.
 FT DOMAIN 2069 2182 KRINGLE TYPE IV, 19.
 FT DOMAIN 2183 2296 KRINGLE TYPE IV, 20.
 FT DOMAIN 2297 2410 KRINGLE TYPE IV, 21.
 FT DOMAIN 2411 2524 KRINGLE TYPE IV, 22.
 FT DOMAIN 2525 2638 KRINGLE TYPE IV, 23.
 FT DOMAIN 2639 2752 KRINGLE TYPE IV, 24.
 FT DOMAIN 2753 2866 KRINGLE TYPE IV, 25.
 FT DOMAIN 2867 2980 KRINGLE TYPE IV, 26.
 FT DOMAIN 2981 3094 KRINGLE TYPE IV, 27.
 FT DOMAIN 3095 3208 KRINGLE TYPE IV, 28.
 FT DOMAIN 3209 3322 KRINGLE TYPE IV, 29.
 FT DOMAIN 3323 3436 KRINGLE TYPE IV, 30.
 FT DOMAIN 3437 3550 KRINGLE TYPE IV, 31.
 FT DOMAIN 3551 3664 KRINGLE TYPE IV, 32.
 FT DOMAIN 3665 3778 KRINGLE TYPE IV, 33.
 FT DOMAIN 3779 3892 KRINGLE TYPE IV, 34.
 FT DOMAIN 3893 3998 KRINGLE TYPE IV, 35.
 FT DOMAIN 3999 4112 KRINGLE TYPE IV, 36.
 FT DOMAIN 4113 4226 KRINGLE TYPE IV, 37.
 FT DOMAIN 4227 4340 KRINGLE TYPE V.
 FT DOMAIN 4341 4454 SERINE PROTEASE.
 FT ACT_SITE 4369 4548 CHARGE RELAY SYSTEM.
 FT ACT_SITE 4412 4412 CHARGE RELAY SYSTEM.
 FT ACT_SITE 4498 4498 CHARGE RELAY SYSTEM.
 FT VARIANT 4193 4193 W -> R (LOSS OF LYSINE-SEPHAROSE BINDING).
 FT FT
 FT SQ SEQUENCE 4548 AA; 501313 MW; 96921BE96A465C5F CRC64;
 Query Match 19.4%; Score 162.5; DB 1; Length 4548;
 Best Local Similarity 33.1%; Pred. No. 1.6e-07;
 Matches 41; Conservative 12; Mismatches 46; Indels 25; Gaps 7;
 QY 27 NIHW--CN---CP-----KFGGHCETDKSKT---CYEGNGHFYRGKASDTDMGR 69
 DB 3742 NVRYEYCNLTQCPVTSESSVLATSTAVSEQAPTEQSPVQDCYHGGQSVRGFSFTVTGR 3801
 QY 70 PCLFWSATVLCQTYHAHRSALQLGLGHNYCRNPDRRRPWCYVQVGLKPLV--QECM 127
 DB 3802 TCQSWSSMT---PHWHQRTTEYPNGGLTRNYCRNPDAEIRPWCYT---MDPSVRWEYCN 3855
 QY 128 VHDC 131
 DB 3856 LTQC 3859
 RESULT 23
 KRM1 RAT
 ID KRM1 RAT STANDARD; PRT; 473 AA.
 AC Q324S4.
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Kremen protein 1 precursor (Kringle-containing protein marking the eye
 DE and the nose) (Dickkopf receptor).
 GN KREMEN1 OR KREMEN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nakamura T., Nakamura T.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
 CC to block Wnt/beta-catenin signaling (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).

```

CC -1- SIMILARITY: Contains 1 CUB domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB065090; BAB2003.1; .
CC GO; GO:0016021; C: integral to membrane; ISS.
CC InterPro; IPR000859; CUB domain.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR002889; WSC.
CC Pfam; PF00431; CUB; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF01822; WSC; 1.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00042; CUB; 1.
CC SMART; SM00130; KR; 1.
CC PROSITE; PS01180; CUB; 1.
CC PROSITE; PS00021; KRINGLE 1; 1.
CC PROSITE; PS50070; KRINGLE 2; 1.
CC Knt signaling pathway; Signal; Transmembrane; Kringle.
FT SIGNAL 1 19
FT CHAIN 20 473
FT KREMER PROTEIN 1.
FT DOMAIN 21 392
FT TRANSMEM 393 413
FT DOMAIN 414 473
FT DOMAIN 31 114
FT DOMAIN 120 210
FT DOMAIN 214 321
FT CARBOHYD 59 59
FT CARBOHYD 217 217
FT CARBOHYD 255 255
FT CARBOHYD 293 293
FT CARBOHYD 333 333
FT CARBOHYD 345 345
SQ SEQUENCE 473 AA; 51869 MW; 9B510857DF856F08 CRC64;

Query Match 19.2%; Score 161; DB 1; Length 473;
Best Local Similarity 45.1%; Pred. No. 2.5e-08;
Matches 32; Conservative 7; Mismatches 28; Indels 4; Gaps 2;

QY 47 SKTCYEGNGHFYRGKASTDTM--GRPCLPWSATVLQQTYYAHRSDALQLGLGKHNVCN 104
Db 29 SPECFTANGADYRGTSWTALQGGKCLFWNE--TFQHPYNTLKYPNGGGGLGHNVCN 86
QY 105 PDNRPRPCVY 115
Db 87 PDGDSVPCVY 97

RESULT 24
APOA MACMU
ID APOA MACMU STANDARD; PRT; 1420 AA.
AC P1417;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Apolipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Lp(a)) (Fragment).
GN LPA.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=89174660; PubMed=2925643;
RA Tomlinson J.E., McLean J.W., Lawn R.M.;
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
RT synthesis."
RL J. Biol. Chem. 264:5957-5965 (1989).
CC -1- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp330.
CC -1- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC decorin (By similarity).
CC -1- PTM: N- and O-glycosylated (By similarity).
CC -1- DISEASE: Elevated plasma concentrations of apo(a) and its
CC naturally occurring proteolytic fragments are correlated with
CC atherosclerosis. Homology with plasminogen kringle IV and V is
CC thought to underlie the atherogenicity of the protein, because the
CC fragments are competing with plasminogen for fibrinogen binding.
CC -1- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
CC leading to the formation of the so called mini-Lp(a). Apo(a)
CC fragments accumulate in atherosclerotic lesions, where they may
CC promote thrombogenesis. O-glycosylation may limit the extent of
CC proteolytic fragmentation (By similarity).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -1- SIMILARITY: Contains at least 10 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J04635; AAA36833.1; .
CC PIR; A32869; A32869.
CC HSSP; P00747.2PK4.
CC MEROPS; S01.226; .
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser protease_Try.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 10.
CC SMART; SM00130; KR; 10.
CC SMART; SM00020; Tryp SPC; 1.
CC PROSITE; PS00021; KRINGLE 1; 10.
CC PROSITE; PS50070; KRINGLE 2; 10.
CC PROSITE; PS50240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; FALSE NEG.
CC PROSITE; PS00135; TRYPSIN_SER; FALSE NEG.
KW Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
KW Kringle; Repeat; Atherosclerosis.
FT NON_TER 1 1
FT DOMAIN 49 127 KRINGLE 1.
FT DOMAIN 163 241 KRINGLE 2.
FT DOMAIN 277 355 KRINGLE 3.
FT DOMAIN 391 469 KRINGLE 4.
FT DOMAIN 505 583 KRINGLE 5.
FT DOMAIN 619 697 KRINGLE 6.
FT DOMAIN 725 803 KRINGLE 7.
FT DOMAIN 839 917 KRINGLE 8.
FT DOMAIN 953 1031 KRINGLE 9.
FT DOMAIN 1067 1145 KRINGLE 10.
FT DOMAIN 1191 1420 SERINE PROTEASE.
SQ SEQUENCE 1420 AA; 158367 MW; BE102949E03C5B0E CRC64;

Query Match 19.2%; Score 161; DB 1; Length 1420;
Best Local Similarity 40.9%; Pred. No. 7.2e-08;
Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;

QY 50 CYEGNGHFYRGKASTDTMGRPCLPWSATVLQ--QTYAHRSDALQLGLGKHNVCNPDN 107

```

```
Db 1068 CYHNGQSVRGTFSTVGTGTCQSSMTPPHQKTPENHPNDLLTM-----NYCRNEDA 1122
QY 108 RRRPCYVQGLKPLVQE--CMWHOCAD 133
Db 1123 DTGPWCFT---MDPSVVRREYCNLTRCSD 1147

RESULT 25
PLMN MACMU
ID PLMN MACMU STANDARD; PRT; 810 AA.
AC P12545;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7).
GN PLG.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]_SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RX MEDLINE=89174660; PubMed=2925643;
RA Tomlinson J.E., McLean J.W., Lawn R.M.;
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
RT synthesis."
RL J. Biol. Chem. 264:5957-5965(1989).
CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC GRAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. ACTIVATED WITH CATALYTIC AMOUNTS OF STREPTOKINASE.
CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -!- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION
CC INVOLVES ONLY CLEAVAGE AFTER ARG-580, RESULTING IN 2 CHAINS HELD
CC TOGETHER BY 2 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE
CC ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -!- SIMILARITY: Contains 5 kringle domains.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J04697; AAA36901.1; -
CC PIR; B32869; B30848.
CC HSSP; P00747; 1PMK.
CC MEROPS; S01.233; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR003014; PAN.
CC InterPro; IPR003609; Pan app.
CC InterPro; IPR003966; Prothrombin.
CC InterPro; IPR001254; Ser protease_Try.
CC Pfam; PF00051; kringle; 5.
CC Pfam; PF00024; PAN; 1.
CC Pfam; PF00089; trypsin; 1.
```

```
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR PRODOM; PD000395; Kringle; 5.
DR SMART; SMO0130; KR; 4.
DR SMART; SMO0473; PAN AP; 1.
DR SMART; SMO0020; Tryp SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
KW Signal.
FT SIGNAL 1 19
FT CHAIN 20 810 PLASMINOGEN.
FT CHAIN 20 580 PLASMIN HEAVY CHAIN A.
FT PEPTIDE 20 96 ACTIVATION PEPTIDE.
FT CHAIN 97 580 PLASMIN SHORT FORM OF CHAIN A.
FT CHAIN 581 810 PLASMIN LIGHT CHAIN B.
FT DOMAIN 103 181 KRINGLE 1.
FT DOMAIN 184 262 KRINGLE 2.
FT DOMAIN 275 352 KRINGLE 3.
FT DOMAIN 377 454 KRINGLE 4.
FT DOMAIN 481 560 KRINGLE 5.
FT ACT_SITE 622 622 SERINE PROTEASE.
FT ACT_SITE 665 665 CHARGE RELAY SYSTEM.
FT ACT_SITE 760 760 CHARGE RELAY SYSTEM.
FT BINDING 136 136 OMEGA-AMINOCARBOXYLIC ACIDS.
FT BINDING 158 158 OMEGA-AMINOCARBOXYLIC ACIDS.
FT BINDING 172 172 OMEGA-AMINOCARBOXYLIC ACIDS.
FT BINDING 432 432 OMEGA-AMINOCARBOXYLIC ACIDS.
FT BINDING 445 445 OMEGA-AMINOCARBOXYLIC ACIDS.
FT BINDING 134 134 FIBRIN.
FT BINDING 136 136 FIBRIN.
FT DISULFID 49 73 BY SIMILARITY.
FT DISULFID 53 61 BY SIMILARITY.
FT DISULFID 103 181 BY SIMILARITY.
FT DISULFID 124 164 BY SIMILARITY.
FT DISULFID 152 176 BY SIMILARITY.
FT DISULFID 185 262 BY SIMILARITY.
FT DISULFID 188 316 BY SIMILARITY.
FT DISULFID 206 245 BY SIMILARITY.
FT DISULFID 234 257 BY SIMILARITY.
FT DISULFID 275 352 BY SIMILARITY.
FT DISULFID 296 335 BY SIMILARITY.
FT DISULFID 324 347 BY SIMILARITY.
FT DISULFID 377 454 BY SIMILARITY.
FT DISULFID 398 437 BY SIMILARITY.
FT DISULFID 426 449 BY SIMILARITY.
FT DISULFID 481 560 BY SIMILARITY.
FT DISULFID 502 543 BY SIMILARITY.
FT DISULFID 531 555 BY SIMILARITY.
FT DISULFID 567 685 BY SIMILARITY.
FT DISULFID 577 585 BY SIMILARITY.
FT DISULFID 607 623 BY SIMILARITY.
FT DISULFID 699 766 BY SIMILARITY.
FT DISULFID 729 745 BY SIMILARITY.
FT DISULFID 756 784 BY SIMILARITY.
FT CARBOHYD 365 365 O-LINKED (GALNAc...) (BY SIMILARITY).
SQ SEQUENCE 810 AA; 90255 MW; A75E1C51A1A0F24A CRC64;
```

Query Match 19.1%; Score 159.5; DB 1; Length 810;
Best Local Similarity 31.4%; Pred. No. 5.8e-08;
Matches 49; Conservative 12; Mismatches 66; Indels 29; Gaps 7;

QY 10 NCDCINGGTCVSNKYFSNIHMCNCKPKKFGGQ-----HCEIDSKSKTCYEGNGH 56

Db 428 NPADKGPWCFTTDPDSVRWEYCNLKKCSGTGSGVAAPPVPAQLPDAETPSEEDCMFGNGK 487

QY 57 FYRGKASTDTMGRPCLPFWNSATVLQOTYHAHR-----SDALQLGLGKHNYCRNPD-NRRRP 111

```

Db 488 GYRGKATVTGTPCEWAA-----GEPHSRIFTBETHPRAGLEK-NYCRAPDGDVGGP 541
QY 112 WCYVQGLKFLVQECMVHOC-----DGKXPSSPPEE 143
Db 542 WCYT-TNPKLFYCDVPQCAASSFDCGKQVPEPKK 576

RESULT 26
KRM1_XENLA STANDARD; PRT; 452 AA.
AC Q90Y30;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringling-containing protein marking the eye
and the nose) (Dickkopf receptor).
GN KREMEN1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Nakamura T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 Kringling domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB070851; BAB64294.1; -.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR000001; Kringling.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringling; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PD00018; KRINGLE.
DR ProDom; PD000395; Kringling; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00321; WSC; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
KW Wnt signaling pathway; Glycoprotein; Kringling;
KW Transmembrane; Signal;
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 452 KREMEN PROTEIN 1.
FT DOMAIN 23 369 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 370 390 POTENTIAL.
FT DOMAIN 391 452 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 29 112 KRINGLE.
FT DOMAIN 113 208 WSC.
FT DOMAIN 212 319 CUB.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 452 AA; 50188 MW; ED24BCDIAF4564E2 CRC64;

Query Match
Best Local Similarity 19.0%; Score 159; DB 1; Length 452;
Matches 31; Conservative 9; Mismatches 32; Indels 4; Gaps 2;

QY 42 CEIDKSKTCYEGNGHFYRGKASTDTM--GRPLPWSATVLQOQTYHAHRSALQGLGCKH 99
DB 22 CSDSFHSECYTVNGADYRGTONQTSLDGKPCLEFNE--TFQHPYNTLKYPNGEGGLGEH 79
QY 100 NYCNPDPNRRPWCYV 115
DB 80 NYCNPDPGDVSPWCYI 95

RESULT 27
KRM1_MOUSE STANDARD; PRT; 473 AA.
ID KRM1_MOUSE
AC Q99N43;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringling-containing protein marking the eye
and the nose) (Dickkopf receptor).
GN KREMEN1 OR KREMEN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.; DEVELOPMENTAL STAGE, AND TISSUE SPECIFICITY.
RC TISSUE=Brain, and Kidney;
RA MEDLINE=21167372; PubMed=11267660;
RA Nakamura T.; Aoki S.; Kitajima K.; Takahashi T.; Matsumoto K.;
RT "Molecular cloning and characterization of Kremen, a novel
Kringling-containing transmembrane protein.";
RL Biochim. Biophys. Acta 1518:63-72(2001).
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: In the adult, widely expressed with high
levels in heart, lung, kidney, skeletal muscle and testis.
CC -!- DEVELOPMENTAL STAGE: In the embryo, expression is first detected
on day 9 and increases up to day 18. Lower levels are found in
adult. At 9.5 dpc, expression is localised to the apical
ectodermal ridge (AER) of the developing fore- and hindlimb buds,
the telencephalon and the first brachial arch. At 10.5 dpc,
expression is also observed in the myotome and in sensory tissues
such as the nasal pit and optic vesicle.
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 Kringling domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB059617; BAB40968.1; -.
DR HSSP; P00747; ICEA.
DR MGD; MGI:193398; Kremen.
DR GO; GO:0016021; C:integral to membrane; NAS.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR000001; Kringling.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringling; 1.
DR Pfam; PF01822; WSC; 1.

```


DR PRINTS; PRO0018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00042; CUB; 1.
 DR SMART; SM00130; KR; 1.
 DR PROSITE; PS01180; CUB; 1.
 DR PROSITE; PS00021; KRINGLE 1; 1.
 DR PROSITE; PS00070; KRINGLE 2; 1.
 KW Wnt signaling pathway; Signal; Transmembrane; Kringle.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 473 KREMEN PROTEIN 1.
 FT DOMAIN 21 392 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 393 433 POTENTIAL.
 FT DOMAIN 414 473 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 31 114 KRINGLE.
 FT DOMAIN 120 210 WSC.
 FT DOMAIN 214 321 CUB.
 FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 473 AA; 51716 MW; 586827788BE3PDD1 CRC64;
 Query Match 19.0%; Score 159; DB 1; Length 473;
 Best Local Similarity 45.6%; Pred. No. 3.8e-08;
 Matches 31; Conservative 7; Mismatches 26; Indels 4; Gaps 2;
 QY 50 CYEGNGHFYRGKASTDTM--GRPCLPWNSATVLQQTYYHAHRSDALQLGLGKHNKRNPDN 107
 Db 32 CFTANGADYRGQTSMALQGGKPCFLWNE--TFQHPYNTLKYPNEGGGLGEHNYCNPDDG 89
 QY 108 RRRPQCYV 115
 Db 90 DVSPMCYV 97
 RESULT 28
 PLMN BOVIN
 ID PLMN BOVIN STANDARD; PRT; 812 AA.
 AC P06868; Q28162;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasminogen precursor (EC 3.4.21.7).
 GN PLG.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]_TaxID=9913;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Berglund L., Andersen M.D., Petersen T.E.;
 RA "Cloning and characterization of the bovine plasminogen cDNA.";
 RL Int. Dairy J. 5:593-603(1995).
 RL [2]
 RP SEQUENCE OF 27-812, AND CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=85203906; PubMed=3846532;
 RA Schaller J., Moser P.W., Dannerger-Muller G.A.K., Rosselet S.J.,
 RA Kampfer U., Rickli E.E.;
 RT "Complete amino acid sequence of bovine plasminogen. Comparison with
 RT human plasminogen.";
 RL Eur. J. Biochem. 149:267-278(1985).
 RL [3]
 RP SEQUENCE OF 706-812 FROM N.A.
 RX MEDLINE=85023311; PubMed=6148961;
 RA Malinowski D.P., Sadler J.E., Davie E.W.;
 RT "Characterization of a complementary deoxyribonucleic acid coding for
 RT human and bovine plasminogen.";
 RL Biochemistry 23:4243-4250(1984).
 RL [4]

RP CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=88185329; PubMed=3356193;
 RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
 RA Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
 RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
 RT plasminogen. Species specificity in relation to sialylation and
 RT fucoylation patterns.";
 RL Eur. J. Biochem. 173:57-63(1988).
 CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
 CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
 CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
 CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
 CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
 CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
 CC LAMININ AND VON WILLEBRAND FACTOR.
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 CC higher selectivity than trypsin. Converts fibrin into soluble
 CC products.
 CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
 CC ACTIVATORS. BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
 CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
 CC -!- PTM: N-LINKED GLYCAN CONTAIN N-ACETYLGLUCOSAMINE AND SIALIC ACID.
 CC O-LINKED GLYCANS CONSIST OF GAL-GALNAc DISACCHARIDE WITH IS
 CC MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES (MICROHETEROGENEITY).
 CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
 CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -!- SIMILARITY: Contains 5 kringle domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X79402; CAA55939.1; -;
 DR EMBL; K02935; AAA30714.1; -;
 DR PIR; S45046; PLBO.
 DR HSSP; P00747; 2PK4.
 DR MEROPS; S01_233; -;
 DR GlycoSuiteDB; P06868; -;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan app.
 DR InterPro; IPR003966; Prothrombin.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00051; Kringle; 5.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR PRINTS; PR01505; PROTHROMBIN.
 DR ProDom; PD000395; Kringle; 5.
 DR SMART; SM00130; KR; 5.
 DR SMART; SM00473; PAN AP; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00021; KRINGLE 1; 5.
 DR PROSITE; PS00070; KRINGLE 2; 5.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
 KW Signal.
 FT SIGNAL
 FT CHAIN 1 26 PLASMINOGEN.
 FT CHAIN 27 812 PLASMIN HEAVY CHAIN A.
 FT CHAIN 584 812 PLASMIN LIGHT CHAIN B.
 FT CHAIN 110 188 KRINGLE 1.
 FT DOMAIN

```
PT DOMAIN 192 269 KRINGLE 2.
PT DOMAIN 282 359 KRINGLE 3.
PT DOMAIN 384 461 KRINGLE 4.
PT DOMAIN 485 564 KRINGLE 5.
PT DOMAIN 584 812 SERINE PROTEASE.
PT CARBOHYD 315 315 N-LINKED (GLCNAC. . .).
PT CARBOHYD 315 315 /FTIG=CAR 000014.
PT CARBOHYD 365 365 O-LINKED (GALNAC. . .).
PT ACT_SITE 624 624 /FTIG=CAR 000015.
PT ACT_SITE 667 667 CHARGE RELAY SYSTEM.
PT ACT_SITE 762 762 CHARGE RELAY SYSTEM.
PT CONFLICT 335 335 N -> D (IN REF. 2).
PT CONFLICT 516 516 Q -> H (IN REF. 2).
PT CONFLICT 555 555 P -> L (IN REF. 2).
PT CONFLICT 744 744 T -> R (IN REF. 3).
SQ SEQUENCE 812-AA; 91216 MW; 38A6AA691E220946 CRC64;

Query Match 19.0%; Score 159; DB 1; Length 812;
Best Local Similarity 28.0%; Pred. No. 6.5e-08;
Matches 49; Conservative 15; Mismatches 55; Indels 56; Gaps 11;

QY 5 HOVPSNCDCLN-----GGTCVSNKYVFSNIHW--CNCBKFKGGQHCIEDKSKT 49
DB 315 NRTPENFPCKLEENYCRNPNNGKAPWCYTN--SEVRWEYCTIPS-----CESSPLST 366
QY 50 -----CYEGNGHYFVGKASTDTMGRCPLPMNSATVLOQTVAH---RS 89
DB 367 ERMDVPVPEQTVPQDCYHNGSQSVRGTSITITGRKQSWSS-----MTPRHLKTPE 421
QY 90 DALQLGLGKHENCRNPNRRPNCYQVGLKPLV--QECMVHDCADGKK--PSSP 140
DB 422 NYENAGL-TMNYCRNPADKSPWCYT---TDPVRWEFCMLKFKCSETPEQVPAAP 472

RESULT 29
ROR1 HUMAN
ID ROR1 HUMAN STANDARD; PRT; 937 AA.
AC Q1973; Q92776; PRT;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase transmembrane receptor ROR1 precursor
DE (SC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 1).
GN ROR1 OR NTRK1
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=93100347; PubMed=1334494;
RA Maslakowski P, Carroll R.D.;
RT "A novel family of cell surface receptors with tyrosine kinase-like
RT domain.";
RL J. Biol. Chem. 267:26181-26190(1992).
RN [2]
SEQUENCE FROM N.A. (ISOFORM SHORT).
RX MEDLINE=97030043; PubMed=8875995;
RA Reddy U.R., Phatak S., Pleasure D.;
RT "Human neural tissues express a truncated Ror1 receptor tyrosine
RT kinase, lacking both extracellular and transmembrane domains.";
RL Oncogene 13:1555-1559(1996).
CC -1- FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHOSE ROLE IS NOT YET
CC CLEAR.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q01973-1; Sequence=Displayed;
CC Name=Short; Synonyms=T-ROR1;
```

```
CC IsoId=Q01973-2; Sequence=VSP_005008;
CC -1- TISSUE SPECIFICITY: EXPRESSED STRONGLY IN HUMAN HEART, LUNG, AND
CC KIDNEY, BUT WEAKLY IN THE CNS. THE SHORT ISOFORM IS STRONGLY
CC EXPRESSED IN FETAL AND ADULT CNS AND IN A VARIETY OF HUMAN
CC CANCERS, INCLUDING THOSE ORIGINATING FROM CNS OR PNS
CC NEUROECTODERM.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS DURING EARLY
CC EMBRYONIC DEVELOPMENT. THE EXPRESSION LEVELS DROP STRONGLY AROUND
CC DAY 16 AND THERE ARE ONLY VERY LOW LEVELS IN ADULT TISSUES. ROR
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ROR
CC SUBFAMILY.
CC -1- SIMILARITY: Contains 1 frizzled (FZ) domain.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M97675; AAA60275.1; -;
CC EMBL; U38894; AAC50714.1; -;
CC PIR; A45082; A45082.
CC HSSP; P00747; 1CEA.
CC Genew; HGNC:10256; ROR1.
CC MIM; 602336; -;
CC GO; GO:0005737; Cytoplasm; TAS.
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0004714; P:transmembrane receptor protein tyrosine kin. ; TAS.
CC GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. ; TAS.
CC InterPro; IPR000024; Fz domain.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_C2.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR00719; Prot_kinase.
CC InterPro; IPR001245; Tyr_kinase.
CC Pfam; PF01392; Fz; 1.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00018; KRINGLE.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000395; Kringle; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00408; IGC2; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00219; TyRK; 1.
CC PROSITE; PS50038; FZ; 1.
CC PROSITE; PS50835; IG_LIKE; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS50070; KRINGLE_2; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC Transferrase; Tyrosine-protein kinase; ATP-binding; Receptor;
CC Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
CC Immunoglobulin domain; Alternative splicing.
CC SIGNAL 1 29
CC CHAIN 30 937
CC TYROSINE-PROTEIN KINASE TRANSMEMBRANE
CC RECEPTOR ROR1.
CC EXTRACELLULAR (POTENTIAL).
CC POTENTIAL.
CC CYTOPLASMIC (POTENTIAL).
CC IG-LIKE C2-TYPE.
CC FZ.
CC KRINGLE.
CC PROTEIN KINASE.
CC SER/THR-RICH.
CC PRO-RICH.
CC 784 851
```

```

FT DOMAIN 853 876 SER/THR-RICH.
FT NP_BIND 479 487 ATP (BY SIMILARITY).
FT BINDING 506 506 ATP (BY SIMILARITY).
FT ACT_SITE 615 615 BY SIMILARITY.
FT MOD_RES 645 645 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT DISULFID 79 131 BY SIMILARITY.
FT CARBOHYD 47 47 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPIC 1 549 Missing (in isoform Short).
FT SEQUENCE 937 AA; 104312 MW; 0D0694DBF29F4773 CRC64;

Query Match 18.9%; Score 158.5; DB 1; Length 937;
Best Local Similarity 29.7%; Pred. No. 8.4e-08;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;

QY 2 NELHQVP-----SNCDCLNGGTCVSNKYFS-----NIHWNC-----PKKFGQH 41
DB 236 DETSSVFKPRDLCDCEILENVLCQTEYIFARSNPMLMLRLKLENCELDLPQSPPEAAN 295
QY 42 C-----EIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSAL 92
DB 296 CIRIGIPWADPINKHKYCNSTGVYRGTVSVTKSGRCQCPWNS-----QYPHTHTFTAL 350
QY 93 QLQ--LGKHNCRNPNRR--PWCY 114
DB 351 RPELNGGHSHYCRNPGNQKEAPWCF 375

RESULT 30
RORL_MOUSE STANDARD; PRT; 937 AA.
AC Q92139;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase transmembrane receptor RORL precursor
DE (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 1)
DE (mRORL).
GN ROR1 OR NTRK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9248426; PubMed=10231392;
RA Oishi I., Takeuchi S., Hashimoto R., Nagabukuro A., Ueda T., Liu Z.J.,
RA Hattori T., Akita S., Matsuda Y., Yamamura H., Otani H., Minami Y.;
RT "Spatio-temporally regulated expression of receptor tyrosine kinases,
RT mRor1, mRor2, during mouse development: implications in development
RT and function of the nervous system.";
RL Genes Cells 4:41-56(1999).
CC -!- FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHOSE ROLE IS NOT YET
CC CLEAR.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ROR
CC SUBFAMILY.
CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; AB010383; BAA75480.1; --
DR HSRP; P00747; ICEA.
DR MGD; MGI:1347520; Ror1.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_C2.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD0000395; Kringle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS50038; FZ; 1.
DR PROSITE; PS50835; IG-Like; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Receptor;
KW Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
KW Immunoglobulin domain.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 937 TYROSINE-PROTEIN KINASE TRANSMEMBRANE
RECEPTOR ROR1.
FT DOMAIN 30 406 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 407 427 POTENTIAL.
FT DOMAIN 428 937 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 42 141 IG-LIKE C2-TYPE.
FT DOMAIN 165 299 FZ.
FT DOMAIN 312 391 KRINGLE.
FT DOMAIN 473 746 PROTEIN KINASE.
FT DOMAIN 753 782 SER/THR-RICH.
FT DOMAIN 784 851 PRO-RICH.
FT DOMAIN 853 876 SER/THR-RICH.
FT NP_BIND 479 487 ATP (BY SIMILARITY).
FT BINDING 506 506 ATP (BY SIMILARITY).
FT ACT_SITE 615 615 BY SIMILARITY.
FT MOD_RES 645 645 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT DISULFID 79 131 BY SIMILARITY.
FT CARBOHYD 47 47 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 937 AA; 104156 MW; D728733E67D1782C CRC64;

Query Match 18.9%; Score 158.5; DB 1; Length 937;
Best Local Similarity 29.7%; Pred. No. 8.4e-08;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;

QY 2 NELHQVP-----SNCDCLNGGTCVSNKYFS-----NIHWNC-----PKKFGQH 41
DB 236 DETSSVFKPRDLCDCEILENVLCQTEYIFARSNPMLMLRLKLENCELDLPQSPPEAAN 295
QY 42 C-----EIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSAL 92
DB 296 CIRIGIPWADPINKHKYCNSTGVYRGTVSVTKSGRCQCPWNS-----QYPHTHTFTAL 350
QY 93 QLQ--LGKHNCRNPNRR--PWCY 114
DB 351 RPELNGGHSHYCRNPGNQKEAPWCF 375

```

Search completed: December 3, 2003, 14:40:10
Job time : 5.59483 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:02 : Search time 22.5259 Seconds
(without alignments)
1007.637 Million cell updates/sec

Title: US-09-880-503-8
Perfect score: 837
Sequence: 1 SNEHLHQVPSNCDLNGTGV.....QECMVHDCADGKXFPSPPEE 143

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_13Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	837	100.0	143	AAE16549	Human uPA amino te
2	837	100.0	337	AAE16549	Human colon cancer
3	837	100.0	337	AAE16549	Human ovarian anti
4	837	100.0	411	AAE16549	Sequence encoded b
5	837	100.0	411	AAE16549	Human native prou
6	837	100.0	411	AAE16549	UK-S3 as encoded b
7	837	100.0	411	AAE16549	Urokinase precurs
8	837	100.0	411	AAE16549	Pro-urokinase deri
9	837	100.0	411	AAE16549	Pro-urokinase deri

10	837	100.0	411	AAE16549	Pro-urokinase. Ho
11	837	100.0	411	AAE16549	Pro-urokinase muta
12	837	100.0	411	AAE16549	Pro-urokinase muta
13	837	100.0	411	AAE16549	Pro-urokinase muta
14	837	100.0	411	AAE16549	Pro-urokinase muta
15	837	100.0	411	AAE16549	Pro-urokinase muta
16	837	100.0	411	AAE16549	Pro-urokinase muta
17	837	100.0	411	AAE16549	Pro-urokinase muta
18	837	100.0	411	AAE16549	Pro-urokinase muta
19	837	100.0	411	AAE16549	Pro-urokinase muta
20	837	100.0	411	AAE16549	Pro-urokinase muta
21	837	100.0	411	AAE16549	Pro-urokinase muta
22	837	100.0	411	AAE16549	Pro-urokinase muta
23	837	100.0	411	AAE16549	Pro-urokinase muta
24	837	100.0	411	AAE16549	Pro-urokinase muta
25	837	100.0	411	AAE16549	Pro-urokinase muta
26	837	100.0	411	AAE16549	Pro-urokinase muta
27	837	100.0	411	AAE16549	Pro-urokinase muta
28	837	100.0	411	AAE16549	Pro-urokinase muta
29	837	100.0	411	AAE16549	Pro-urokinase muta
30	837	100.0	411	AAE16549	Pro-urokinase muta
31	837	100.0	411	AAE16549	Pro-urokinase muta
32	837	100.0	411	AAE16549	Pro-urokinase muta
33	837	100.0	411	AAE16549	Pro-urokinase muta
34	837	100.0	411	AAE16549	Pro-urokinase muta
35	837	100.0	411	AAE16549	Pro-urokinase muta
36	837	100.0	411	AAE16549	Pro-urokinase muta
37	837	100.0	411	AAE16549	Pro-urokinase muta
38	837	100.0	411	AAE16549	Pro-urokinase muta
39	837	100.0	411	AAE16549	Pro-urokinase muta
40	837	100.0	411	AAE16549	Pro-urokinase muta
41	837	100.0	411	AAE16549	Pro-urokinase muta
42	837	100.0	411	AAE16549	Pro-urokinase muta
43	837	100.0	411	AAE16549	Pro-urokinase muta
44	837	100.0	411	AAE16549	Pro-urokinase muta
45	837	100.0	411	AAE16549	Pro-urokinase muta

ALIGNMENTS

RESULT 1
AAE16549
ID AAE16549 standard; Protein; 143 AA.
AC AAE16549;
XX
XX
DT 09-APR-2002 (first entry)
DE Human uPA amino terminal fragment (ATF) and connecting peptide.
DE Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
clotting disorder; uterine contraction disorder; respiratory disease;
adult respiratory distress syndrome; amino terminal fragment; ATF;
male impotence.
XX
OS Homo sapiens.
XX
XX WO200197752-A2.
XX
XX 27-DEC-2001.
XX
XX 13-JUN-2001; 2001WO-US18976.
XX
XX 20-JUN-2000; 2000US-212874P.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX Cines DB, Higazi AA;
XX

DR WPI; 2002-122240/16.
DR N-PSDB; AAD27052.
XX
PT Composition for modulating muscle cell and tissue contractility for
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
PT comprising domains from urokinase-type plasminogen activator -
XX
XX Claim 24; Fig 1H; 117pp; English.
XX
XX The invention relates to a composition comprising one or more domains of
CC urokinase-type plasminogen activator (uPA). The composition is used to
CC modulate the contractility and angiogenic activity of a mammalian muscle,
CC endothelial cell or tissue. The composition is used for treating stroke,
CC hypertension, hypertension, atherosclerosis, heart attack, microvascular
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
CC diabetic retinopathy, wound healing, clotting disorder, uterine
CC contraction disorder, male impotence, respiratory disease or condition
CC such as asthma, adult respiratory distress syndrome, primary pulmonary
CC hypertension, microvascular thrombotic occlusion, and a disorder
CC associated with chronic intrapulmonary fibrin formation. The present
CC sequence is human urokinase-type plasminogen activator (uPA) amino
CC terminal fragment (ATP) and connecting peptide.
XX
XX Sequence 143 AA;
SQ
Query Match 100.0%; Score 837; DB 23; Length 143;
Best Local Similarity 100.0%; Pred. No. 1.9e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGTGVSNKYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 60
DB 1 SNEHQVPSNCDCLNGTGVSNKYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWSATVLTQTYVHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWSATVLTQTYVHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPSPPPEE 143
DB 121 PLVQECMVHDCADGKPSPPPEE 143
RESULT 2
ID AAG75492 standard; Protein; 337 AA.
XX
AC AAG75492;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:6256.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 10.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX WPI; 2001-235357/24.
XX

DR N-PSDB; AAH34897.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 11; Page 7707-7708; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77789 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell,
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAG77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX Sequence 337 AA;
SQ
Query Match 100.0%; Score 837; DB 22; Length 337;
Best Local Similarity 100.0%; Pred. No. 3.8e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGTGVSNKYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 60
DB 27 SNEHQVPSNCDCLNGTGVSNKYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRG 86
QY 61 KASTDTMGRPCLPWSATVLTQTYVHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLK 120
DB 87 KASTDTMGRPCLPWSATVLTQTYVHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLK 146
QY 121 PLVQECMVHDCADGKPSPPPEE 143
DB 147 PLVQECMVHDCADGKPSPPPEE 169
RESULT 3
ID ABP41795 standard; Protein; 337 AA.
XX
AC ABP41795;
XX
DT 22-AUG-2002 (first entry)
XX
DE Human ovarian antigen HVCEB79, SEQ ID NO:2927.
XX
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
XX inflammatory condition; immune disorder; blood disorder;
XX cardiovascular disorder; respiratory disorder; neurological disorder;
XX gastrointestinal disorder; urinary system disorder; drug screening;
XX gene therapy; chromosome mapping; forensic analysis;
XX antibody preparation; cytostatic; immunomodulatory; neuroprotective;
XX antiinflammatory; gynaecological; reproductive; chromosome 10q24.
XX
OS Homo sapiens.
XX
PN WO200200677-A1.
XX
XX 03-JAN-2002.
XX

```

PF 07-JUN-2001; 2001WO-US18569.
XX
PR 07-JUN-2000; 2000US-209467P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
XX
XX WPI; 2002-147878/19.
XX N-PSDB; ABQ54872.
XX
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX useful in the prevention, treatment and diagnosis of cancer (e.g.
XX ovarian cancer), immune disorders, cardiovascular disorders and
XX neurological diseases.
XX
XX Claim 11; SEQ ID NO 2927; 2922pp; English.
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prognosing or preventing various ovary and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumours of ovarian or breast origin, reproductive system
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX vaginitis), immune disorders (e.g., congenital and acquired
XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX blood-related disorders (e.g., anemia), cardiovascular disorders,
XX respiratory disorders, neurological disorders, gastrointestinal disorders
XX and urinary system disorders. Ovarian antigen polypeptides and
XX polynucleotides may also be used in screening for compounds which
XX modulate ovarian antigen expression or activity. The polynucleotides may
XX further be used for gene therapy, chromosome mapping, in the
XX identification of individuals and in forensic analysis, and the
XX polypeptides may be used as food additives or to prepare antibodies
XX useful in disease diagnosis, drug targeting and phenotyping. The present
XX sequence represents a human ovarian antigen of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 337 AA;
XX
XX Query Match 100.0%; Score 837; DB 23; Length 337;
XX Best Local Similarity 100.0%; Pred. No. 3.8e-54;
XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHYRG 60
DB 27 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHYRG 86
QY 61 KASDTDMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 87 KASDTDMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 146
QY 121 PLVQECMWHDCAADGKKPSSPPEE 143
DB 147 PLVQECMWHDCAADGKKPSSPPEE 169
XX
RESULT 4
AAP50871
ID AAP50871 standard; protein; 411 AA.
XX
XX AAP50871;
XX

```

```

DT 30-NOV-1991 (first entry)
XX
DE Sequence encoded by cDNA sequence for human urokinase zymogen
DE (Japanese Patent Application No.37119/84).
XX
XX Thrombolytic agent; plasminogen activator activity; fibrin affinity;
XX enzyme.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Cleavage-site 158..159
XX /note= "potential cleavage site which generates
XX the two-chain form from the zymogen"
XX
XX Disulfide-bond 50..131
XX Disulfide-bond 71..113
XX Disulfide-bond 102..126
XX Disulfide-bond 148..279
XX Disulfide-bond 189..205
XX Disulfide-bond 197..268
XX Disulfide-bond 293..362
XX Disulfide-bond 325..341
XX Disulfide-bond 352..380
XX
XX EPI39447-A
XX 02-MAY-1985.
XX
XX 07-SEP-1984; 84EP-0306117.
XX
XX 17-OCT-1983; 83JP-0195051.
XX 13-SEP-1983; 83JP-0170354.
XX
XX (GREC ) GREEN CROSS CORP.
XX
XX Kasai S, Arimura H, Mori K, Nishida M, Suyama T;
XX WPI; 1985-106530/18.
XX
XX New urokinase zymogen - useful as thrombolytic agent
XX
XX Disclosure; Page 12; 30pp; English.
XX
XX Zymogen AAP50871 is the inactive precursor form of human urokinase.
XX Urokinase zymogen is cleaved into the two-chain form composed of
XX characteristic urokinase H (molecular wt. of 30,000) and L (molecular
XX wt. of 20,000) chains when treated with catalytic amounts of plasmin.
XX The patentors claim a new urokinase zymogen which has mol. wt. ca.
XX 50,000, a single chain molecular structure, and selective affinity
XX for fibrin. It is a thrombolytic agent which manifests its
XX plasminogen activator activity on cleavage by proteolytic enzymes
XX (e.g. plasmin) and has higher affinity for fibrin than known forms
XX of urokinase.
XX
XX Sequence 411 AA;
XX
XX Query Match 100.0%; Score 837; DB 6; Length 411;
XX Best Local Similarity 100.0%; Pred. No. 4.4e-54;
XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHYRG 60
DB 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHYRG 60
QY 61 KASDTDMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 61 KASDTDMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
QY 121 PLVQECMWHDCAADGKKPSSPPEE 143
DB 121 PLVQECMWHDCAADGKKPSSPPEE 143
XX

```

RESULT 5

AAW13634
ID AAW13634 standard; Protein; 411 AA.
XX
AC AAW13634;
XX
DT 04-JUN-1997 (first entry)
XX
DE Human native prourokinase.
XX
KW Human; prourokinase; hPUK; variant; half-life; increase; EGF;
KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 10..42 EGF domain
FT /label= EGF domain
FT /note= "in claimed variants, at least part of the
FT EGF domain is deleted (see comments)"
FT
FT Region 10..19
FT /label= first_loop
FT Region 20..31
FT /label= second_loop
FT Region 33..42
FT /label= third_loop
XX
EP398361-A.
PN
XX
PD 22-NOV-1990.
XX
PF 18-MAY-1990; 90EP-0109472.
XX
PR 22-FEB-1990; 90JP-0042020.
PR 18-MAY-1989; 89JP-0126433.
PR 03-JUL-1986; 86JP-0156936.
PR 18-FEB-1987; 87JP-0036495.
PR 18-MAY-1989; 89JP-0126434.
XX
PA (GRC) GREEN CROSS CORP.
XX
PI Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;
PI Morita M, Tanabe T;
XX
DR WPI; 1990-350146/47.
DR N-PSDB; AAI61671.
XX
PT Human pro-urokinase variants - deficient in loop regions of
PT epidermal growth factor, showing long blood half-life, as
PT fibrinolytic agent
XX
PS Disclosure; Fig 1; 22pp; English.
XX
CC New variants of human prourokinase (hPUK) comprise a hPUK deficient
CC in (i) at least part of the first loop region of the epidermal growth
CC factor (EGF) domain; (ii) at least part of the first loop and at
CC least part of the second loop; or (iii) at least part of the third
CC loop. The hPUK variants show an increased blood half-life comparable
CC to that of the whole EGF domain-deficient hPUK variant and urokinase
CC while retaining the same properties as those of hPUK. They have
CC potent thrombolytic activity and very little tendency to cause
CC spontaneous bleeding. The present sequence is that of the wild-type
CC hPUK protein, including the EGF domain.
XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 837; DB 11; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHEIDKSKTCYEGNGHFRG 60
DB 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHEIDKSKTCYEGNGHFRG 60

QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRRRPWCYVQVGLK 120
DB |||||
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRRRPWCYVQVGLK 120
DB |||||
QY 121 PLVQECWVHDCADGKFPSPPEE 143
DB |||||
QY 121 PLVQECWVHDCADGKFPSPPEE 143
DB |||||

RESULT 6

AAAR05117
ID AAAR05117 standard; protein; 411 AA.
XX
AC AAAR05117;
XX
DT 25-MAR-2003 (updated)
DT 04-OCT-1990 (first entry)
XX
DE UK-S3 as encoded by PUKS3.
XX
KW Urokinase; glycosylation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_difference 153
FT /label=synthetic_mutation
FT /note="old seq (leu)"
FT misc_difference 155
FT /label=synthetic_mutation
FT /note="old seq (Pro)"
XX
PN EP370205-A.
XX
PD 30-MAY-1990.
XX
PF 28-SEP-1989; 89EP-0117981.
XX
PR 29-SEP-1988; 88JP-0245705.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Sasaki K, Nishi T, Yasumura S, Sato M, Itoh S;
XX
DR WPI; 1990-165029/22.
DR N-PSDB; AAQ04486.
XX
PT Polypeptide(s) with added carbohydrate chains - formed by
PT modification of amino acid sequence, used to improve
PT physiochemical properties and/or activities.
XX
PS Disclosure; Page ?; 30pp; English.
XX
CC The polypeptide is a deriv. of mature urokinase, designated
CC UK-S3 which has 2 amino acid substituents, which result in an N-linked
CC glycosylation site giving the new protein improved stability and
CC activity.
CC See also AAAR05113-17.
CC (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 837; DB 11; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHEIDKSKTCYEGNGHFRG 60
DB 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHEIDKSKTCYEGNGHFRG 60

QY 61 KASTDTMGRPCLPWNATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
AC
DB 61 KASTDTMGRPCLPWNATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 PLVQECMVHDCADGKKPSPPEE 143
RESULT 7
AAR06244
ID AAR06244 standard; protein; 411 AA.
XX
AC AAR06244;
XX
DT 07-DEC-1990 (first entry)
XX
DE Urokinase precursor protein.
XX
KW Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis;
KW myocardial infarction;
XX
OS Homo sapiens.
XX
PN EP380334-A.
XX
PD 01-AUG-1990.
XX
PF 25-JAN-1990; 90EP-0300772.
XX
PR 17-MAY-1989; 89JP-0121405.
PR 27-JAN-1989; 89JP-0016406.
XX
PA (GREC) GREEN CROSS CORP.
XX
PI Matsuda H, Ueda Y, Tamaouchi K;
XX
DR WPI; 1990-233117/31.
XX
PT Urokinase precursor-lipid composite - used as thrombolytic agent,
PT having prolonged half-life in the blood, enhanced
PT bio-availability and improved activity
XX
PS Claim 3; Fig 1; 11pp; English.
XX
CC By forming a precursor-lipid composite, the half-life of this
CC thrombolytic agent in the blood may be increased, exhibiting
CC improved activity without abnormal acceleration of fibrinolytic
CC activity. Compound is useful as a thrombolytic agent in
CC treatment of cerebral thrombosis, myocardial infarction etc.
XX
SQ Sequence 411 AA;
Query Match 100.0%; Score 837; DB 11; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNEHQVPSNCDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 PLVQECMVHDCADGKKPSPPEE 143
RESULT 8
AAR10057
ID AAR10057 standard; protein; 411 AA.
XX

XX AAR10057;
XX 18-MAR-1991 (first entry)
DT Pro-urokinase derivative UK-T4 with Ser(153) and Thr(155).
DE pro-urokinase; UK-T4; plasminogen activator; myocardial infarction;
KW cerebral thrombosis.
XX
OS Homo sapiens.
XX
PN EP405285-A.
XX
PD 02-JAN-1991.
XX
PF 18-JUN-1990; 90EP-0111471.
XX
PR 19-JUN-1989; 89JP-0156302.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Yasamura S, Nishi T, Ito S;
XX
DR WPI; 1991-008678/02.
DR N-PSDS; AAQ10169.
XX
PT New plasminogen activator almost identical to natural
PT pro-urokinase - is thrombin resistant and used for
PT prophylaxis, treatment of cerebral thrombosis or myocardial
PT infarction
XX
PS Disclosure; Page 8; 84pp; English.
XX
CC UK-T4 is one example of a plasminogen activator which differs from
CC natural human pro-urokinase at positions 153 and 155 (Leu
CC substituted by Ser; Pro substituted by Thr, respectively).
CC The derivative has decreased susceptibility to thrombin compared to
CC natural type pro-UK and higher specific activity.
CC See also AAQ10168 and AAQ10170.
XX
SQ Sequence 411 AA;
Query Match 100.0%; Score 837; DB 12; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNEHQVPSNCDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNATVLTQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 PLVQECMVHDCADGKKPSPPEE 143
RESULT 9
AAR10058
ID AAR10058 standard; protein; 411 AA.
XX
AC AAR10058;
XX
DT 18-MAR-1991 (first entry)
XX
DE Pro-urokinase derivative UK-S3 with Asn(153) and Thr(155).
KW pro-urokinase; UK-S3; plasminogen activator; myocardial infarction;
KW cerebral thrombosis.
XX

```

OS Homo sapiens.
XX EP405285-A.
XX PD 02-JAN-1991.
XX PF 18-JUN-1990; 90EP-0111471.
XX PR 19-JUN-1989; 89JP-0156302.
XX PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX PI Yasamura S, Nishi T, Ito S;
XX DR WPI; 1991-008678/02.
XX DR N-PSDB; AAQ10170.
XX PT New plasminogen activator almost identical to natural
XX PT pro-urokinase - is thrombin resistant and used for
XX PT prophylaxis, treatment of cerebral thrombosis or myocardial
XX PT infarction
XX PS Disclosure; Page 9; 84pp; English.
XX CC UK-S3 is one example of a plasminogen activator which differs from
XX CC natural human pro-urokinase at positions 153 and 155. (Leu
XX CC substituted by Asn; Pro substituted by Thr, respectively).
XX CC The derivative has decreased susceptibility to thrombin compared to
XX CC natural type pro-UK and higher specific activity.
XX CC See also AAQ10168 and AAQ10169.
XX SQ Sequence 411 AA;

Query Match 100.0%; Score 837; DB 12; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHGVSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHGVSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRCPLPWNATVLTQTYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 120
DB 61 KASTDTMGRCPLPWNATVLTQTYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 PLVQECMVHDCADGKKPSPPEE 143

RESULT 10
AAR40225
ID AAR40225 standard; Protein; 411 AA.
AC AAR40225;
XX 10-FEB-1994 (first entry)
DE PUK.
KW Pre-urokinase; thrombolytic; blood; plasmid; PUK.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Region 10..60
XX /note= "Ser, Asn, Pro, Gly, or Tyr in the region
XX 10 to 60 in the N-terminal of the human PUK
XX can be replaced by Thr, Pro or Ala"
XX Misc-difference 64
XX /note= "Thr encoded by AGT (sic)"
XX Misc-difference 177
XX /note= "Thr encoded by TAC (sic)"

OS Homo sapiens.
XX Key Location/Qualifiers
XX Region 10..60
XX /note= "Ser, Asn, Pro, Gly, or Tyr in the region
XX 10 to 60 in the N-terminal of the human PUK
XX can be replaced by Thr, Pro or Ala"
XX Misc-difference 64
XX /note= "Thr encoded by AGT (sic)"
XX Misc-difference 177
XX /note= "Thr encoded by TAC (sic)"

FT Domain 10..49
FT /label= EGF
FT Region 10..19
FT /label= Loop_1
FT Region 20..31
FT /label= Loop_2
FT Region 33..42
FT /label= Loop_3
XX JP05192142-A.
XX 03-AUG-1993.
XX 20-JAN-1992; 92JP-0030178.
XX 20-JAN-1992; 92JP-0030178.
XX (GREG ) GREEN CROSS CORP.
XX WPI; 1993-277461/35.
XX N-PSDB; AAQ48228.
XX Mutant human pre-urokinase - by replacing specified aminoacid(s)
XX in N-terminal for providing longer half-life in blood and higher
XX thrombolytic ability
XX Claim 1; Page 14-16; 26pp; Japanese.
XX Human pre-urokinase can be modified by replacing: Ser, Asn,
XX Pro, Gly, or Tyr in the region ranging from residue 10 to 60 in the
XX N-terminal of the human PUK by Thr, Pro or Ala. These mutants have a
XX longer half-life in the blood and a higher thrombolytic ability.
XX For examples see (AAR47956-R47960).
XX SQ Sequence 411 AA;

Query Match 100.0%; Score 837; DB 14; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHGVSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHGVSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRCPLPWNATVLTQTYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 120
DB 61 KASTDTMGRCPLPWNATVLTQTYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 PLVQECMVHDCADGKKPSPPEE 143

RESULT 11
AAR62991
ID AAR62991 standard; protein; 411 AA.
AC AAR62991;
XX 25-MAR-2003 (updated)
XX 21-SEP-1995 (first entry)
XX Pro-urokinase.
XX Pro-urokinase; thrombolysis; fibrin clot lysis.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Domain 297..313
XX /note= "flexible loop"
XX Disulfide-bond 11...19
XX Disulfide-bond 13...31

```

FT	Disulfide-bond	33..42
FT	Disulfide-bond	50..131
FT	Disulfide-bond	71..113
FT	Disulfide-bond	102..126
FT	Disulfide-bond	148..279
FT	Disulfide-bond	189..205
FT	Disulfide-bond	197..268
FT	Disulfide-bond	293..362
FT	Disulfide-bond	325..341
XX	Disulfide-bond	352..380
XX	WO9501427-A1.	
PN		
XX		
PD	12-JAN-1995.	
XX		
PF	28-JUN-1994;	94WO-US02728.
XX		
PR	02-JUL-1993;	93US-0087163.
XX		
PA	(NEW-) NEW ENGLAND DEACONESS HOSPITAL.	
XX		
PI	Gurewich V, Liu J;	
XX		
XX	WPI; 1995-060991/08.	
DR		
XX		
PT	Pro-urokinase mutants - have thrombolytic activity but reduced fibrinogenolysis activity and non-specific plasminogen activation	
XX		
PS	Disclosure; Fig 1; 46pp; English.	
XX		
AA	AAR62991 is the wild type pro-urokinase, from which the new mutants described in AAR62992-R63008 were derived. These mutants retain the thrombolytic activity of the wild type protein, useful for the treatment of thromboembolism, but have a reduced fibrinogenolysis activity and non-specific plasminogen activation. The mutants can therefore be used for the lysis of fibrin clots without inducing systemic bleeding, as can be the case with the wild type protein. (Updated on 25-MAR-2003 to correct PN field.)	
CC		
XX		
SQ	Sequence	411 AA;
	Query Match	100.0%; Score 837; DB 16; Length 411;
	Best Local Similarity	100.0%; Pred. No. 4.4e-54;
	Matches 143; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 SNEHQVFSNCDCLNGTCVSNKYFNSIHWMCNCFKFGGQHCEIDSKTCTEENGHFFRG	60
Db	1 SNEHQVFSNCDCLNGTCVSNKYFNSIHWMCNCFKFGGQHCEIDSKTCTEENGHFFRG	60
QY	61 KASTDTMGRCPLPWN SATVLQOTYHAHRSDALQLGCKHNYCRNPDRRPPWCYVOVG LK	120
Db	61 KASTDTMGRCPLPWN SATVLQOTYHAHRSDALQLGCKHNYCRNPDRRPPWCYVOVG LK	120
QY	121 PLVQECMVHD CADGGKKPSPPEE	143
Db	121 PLVQECMVHD CADGGKKPSPPEE	143
RESULT 12		
AAR62992		
ID	AAR62992 standard; protein; 411 AA.	
XX		
AC	AAR62992;	
XX		
JT	25-MAR-2003 (updated)	
JT	21-SEP-1995 (first entry)	
XX		
DE	Pro-urokinase mutant Ala300.	
XX		
KW	Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Ala300;	
XW	reduced fibrinogenolysis; non-specific plasminogen activation;	
GW	systemic bleeding.	
XX		

XX DE Pro-urokinase mutant His300.
XX AC AAR62994;
XX AC AAR62994;
XX KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant His300;
XX KW reduced fibrinogenolysis; non-specific plasminogen activation;
XX KW systemic bleeding.
XX OS Homo sapiens.
XX EH Key Location/Qualifiers
FT Domain 297..313
FT Disulfide-bond 11..19 /note= "flexible loop"
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO9501427-A1.
XX PD 12-JAN-1995.
XX PF 28-JUN-1994; 94WO-US07278.
XX PF 02-JUL-1993; 93US-0087163.
XX PA (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
XX PI Gurewich V, Liu J;
XX PI WPI; 1995-060991/08.
XX PT Pro-urokinase mutants - have thrombolytic activity but reduced
XX PT fibrinogenolysis activity and non-specific plasminogen activation
XX PS Claim 5; Fig 1; 46pp; English.
XX CC AAR62991 is the wild type pro-urokinase, from which the new mutants
XX CC described in AAR62992-R63008 were derived. These mutants retain the
XX CC thrombolytic activity of the wild type protein, useful for the
XX CC treatment of thromboembolism, but have a reduced fibrinogenolysis
XX CC activity and non-specific plasminogen activation. The mutants can
XX CC therefore be used for the lysis of fibrin clots without inducing
XX CC systemic bleeding, as can be the case with the wild type protein.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 411 AA;
Query Match 100.0%; Score 837; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSDCCLNGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFRG 60
DB 1 SNELHQPNSDCCLNGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFRG 60
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120
DB 61 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120
QY 121 PLVQECMVHDCADGKSPPEE 143
DB 121 PLVQECMVHDCADGKSPPEE 143
RESULT 14

AAR62994
ID AAR62994 standard; protein; 411 AA.
XX AC AAR62994;
XX AC AAR62994;
XX DT 25-MAR-2003 (updated)
XX DT 21-SEP-1995 (first entry)
XX DE Pro-urokinase mutant His299 Ala300.
XX KW Pro-urokinase; thrombolysis; fibrin clot lysis;
XX KW reduced fibrinogenolysis; non-specific plasminogen activation;
XX KW systemic bleeding; mutant His 299 Ala300.
XX OS Homo sapiens.
XX EH Key Location/Qualifiers
FT Domain 297..313 /note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO9501427-A1.
XX PD 12-JAN-1995.
XX PF 28-JUN-1994; 94WO-US07278.
XX PF 02-JUL-1993; 93US-0087163.
XX PA (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
XX PI Gurewich V, Liu J;
XX PI WPI; 1995-060991/08.
XX PT Pro-urokinase mutants - have thrombolytic activity but reduced
XX PT fibrinogenolysis activity and non-specific plasminogen activation
XX PS Claim 7; Fig 1; 46pp; English.
XX CC AAR62991 is the wild type pro-urokinase, from which the new mutants
XX CC described in AAR62992-R63008 were derived. These mutants retain the
XX CC thrombolytic activity of the wild type protein, useful for the
XX CC treatment of thromboembolism, but have a reduced fibrinogenolysis
XX CC activity and non-specific plasminogen activation. The mutants can
XX CC therefore be used for the lysis of fibrin clots without inducing
XX CC systemic bleeding, as can be the case with the wild type protein.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 411 AA;
Query Match 100.0%; Score 837; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSDCCLNGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFRG 60
DB 1 SNELHQPNSDCCLNGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFRG 60
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120
DB 61 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120

QY 121 PLVQECMVHDCADGKPKSPPEE 143
 Db 121 PLVQECMVHDCADGKPKSPPEE 143

RESULT 15
 AAR62995
 ID AAR62995 standard; protein; 411 AA.
 XX AAR62995;
 AC AAR62995;
 XX AAR62995;
 DT 25-MAR-2003 (updated)
 DT 21-SEP-1995 (first entry)
 DE Pro-urokinase mutant Ala300 Ala301.
 XX Pro-urokinase; thrombolysis; fibrin clot lysis;
 KW reduced fibrinogenolysis; non-specific plasminogen activation;
 KW mutant Ala300 Ala301; systemic bleeding.
 OS Homo sapiens.

Key Location/Qualifiers
 Domain 297..313
 Disulfide-bond 11..19
 Disulfide-bond 13..31
 Disulfide-bond 33..42
 Disulfide-bond 50..131
 Disulfide-bond 71..113
 Disulfide-bond 102..126
 Disulfide-bond 148..279
 Disulfide-bond 189..205
 Disulfide-bond 197..268
 Disulfide-bond 293..362
 Disulfide-bond 325..341
 Disulfide-bond 352..380

WO9501427-A1.
 12-JAN-1995.
 28-JUN-1994; 94WO-US07278.
 02-JUL-1993; 93US-0087163.
 (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
 Gurewich V, Liu J;
 WPI; 1995-060991/08.

Pro-urokinase mutants - have thrombolytic activity but reduced
 fibrinogenolysis activity and non-specific plasminogen activation
 Claim 9; Fig 1; 46pp; English.

AAR62991 is the wild type pro-urokinase, from which the new mutants
 described in AAR62992-R63008 were derived. These mutants retain the
 thrombolytic activity of the wild type protein, useful for the
 treatment of thromboembolism, but have a reduced fibrinogenolysis
 activity and non-specific plasminogen activation. The mutants can
 therefore be used for the lysis of fibrin clots without inducing
 systemic bleeding, as can be the case with the wild type protein.
 (Updated on 25-MAR-2003 to correct PN field.)

Query Match 100.0%; Score 837; DB 16; Length 411;
 Best Local Similarity 100.0%; Pred. No. 4.4e-54;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFGGQHCETDKSKTCYEGNGHFYRG 60
 Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHMCNCPKFGGQHCETDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGRPCLPWSATVLQOQTYHAHRS DALQLGLGKHNYCRNPDRRRRRCWYVQVLK 120
 Db 61 KASTDTMGRPCLPWSATVLQOQTYHAHRS DALQLGLGKHNYCRNPDRRRRRCWYVQVLK 120

QY 121 PLVQECMVHDCADGKPKSPPEE 143
 Db 121 PLVQECMVHDCADGKPKSPPEE 143

RESULT 16
 AAR62996
 ID AAR62996 standard; protein; 411 AA.
 XX AAR62996;
 AC AAR62996;
 XX AAR62996;
 DT 25-MAR-2003 (updated)
 DT 21-SEP-1995 (first entry)
 DE Pro-urokinase mutant His300 Ala301.
 XX Pro-urokinase; thrombolysis; fibrin clot lysis;
 KW reduced fibrinogenolysis; non-specific plasminogen activation;
 KW mutant His300 Ala301; systemic bleeding.
 OS Homo sapiens.

Key Location/Qualifiers
 Domain 297..313
 Disulfide-bond 11..19
 Disulfide-bond 13..31
 Disulfide-bond 33..42
 Disulfide-bond 50..131
 Disulfide-bond 71..113
 Disulfide-bond 102..126
 Disulfide-bond 148..279
 Disulfide-bond 189..205
 Disulfide-bond 197..268
 Disulfide-bond 293..362
 Disulfide-bond 325..341
 Disulfide-bond 352..380

WO9501427-A1.
 12-JAN-1995.
 28-JUN-1994; 94WO-US07278.
 02-JUL-1993; 93US-0087163.
 (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
 Gurewich V, Liu J;
 WPI; 1995-060991/08.

Pro-urokinase mutants - have thrombolytic activity but reduced
 fibrinogenolysis activity and non-specific plasminogen activation
 Claim 9; Fig 1; 46pp; English.

AAR62991 is the wild type pro-urokinase, from which the new mutants
 described in AAR62992-R63008 were derived. These mutants retain the
 thrombolytic activity of the wild type protein, useful for the
 treatment of thromboembolism, but have a reduced fibrinogenolysis
 activity and non-specific plasminogen activation. The mutants can
 therefore be used for the lysis of fibrin clots without inducing
 systemic bleeding, as can be the case with the wild type protein.
 (Updated on 25-MAR-2003 to correct PN field.)

Query Match 100.0%; Score 837; DB 16; Length 411;
 Best Local Similarity 100.0%; Pred. No. 4.4e-54;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

XX SQ Sequence 411 AA;
Query Match 100.0%; Score 837; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120

QY 121 PLVQECMVHDCADGKXPSPPEE 143
DB 121 PLVQECMVHDCADGKXPSPPEE 143

RESULT 17
AAR62997
ID AAR62997 standard; protein; 411 AA.
XX AC AAR62997;
XX DT 25-MAR-2003 (updated)
XX DT 21-SEP-1995 (first entry)
XX DE Pro-urokinase mutant Gly306.
XX KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Gly306;
XX KW reduced fibrinogenolysis; non-specific plasminogen activation;
XX KW systemic bleeding.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Domain /note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO9501427-A1.
XX PD 12-JAN-1995.
XX PF 28-JUN-1994; 94WO-US07278.
XX PR 02-JUL-1993; 93US-0087163.
XX PA (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
XX PI Gurewich V, Liu J;
XX DR WPI; 1995-060991/08.
XX KW Pro-urokinase mutants - have thrombolytic activity but reduced
XX KW fibrinogenolysis activity and non-specific plasminogen activation
XX PS Claim 13; Fig 1; 46pp; English.
XX CC AAR62991 is the wild type pro-urokinase, from which the new mutants

```

```

CC described in AAR62992-R63008 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the
CC treatment of thromboembolism, but have a reduced fibrinogenolysis
CC activity and non-specific plasminogen activation. The mutants can
CC therefore be used for the lysis of fibrin clots without inducing
CC systemic bleeding, as can be the case with the wild type protein.
CC (Updated on 25-MAR-2003 to correct FN field.)
XX SQ Sequence 411 AA;
Query Match 100.0%; Score 837; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120

QY 121 PLVQECMVHDCADGKXPSPPEE 143
DB 121 PLVQECMVHDCADGKXPSPPEE 143

RESULT 18
AAR62998
ID AAR62998 standard; protein; 411 AA.
XX AC AAR62998;
XX DT 25-MAR-2003 (updated)
XX DT 21-SEP-1995 (first entry)
XX DE Pro-urokinase mutant Ala313.
XX KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Ala313;
XX KW reduced fibrinogenolysis; non-specific plasminogen activation;
XX KW systemic bleeding.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Domain /note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO9501427-A1.
XX PD 12-JAN-1995.
XX PF 28-JUN-1994; 94WO-US07278.
XX PR 02-JUL-1993; 93US-0087163.
XX PA (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
XX PI Gurewich V, Liu J;
XX DR WPI; 1995-060991/08.

```

XX Pro-urokinase mutants - have thrombolytic activity but reduced
 PT fibrinogenolysis activity and non-specific plasminogen activation
 XX Claim 11; Fig 1; 46pp; English.
 XX AAR62991 is the wild type pro-urokinase, from which the new mutants
 CC described in AAR62992-R63008 were derived. These mutants retain the
 CC thrombolytic activity of the wild type protein, useful for the
 CC treatment of thromboembolism, but have a reduced fibrinogenolysis
 CC activity and non-specific plasminogen activation. The mutants can
 CC therefore be used for the lysis of fibrin clots without inducing
 CC systemic bleeding, as can be the case with the wild type protein.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 411 AA;

Query Match 100.0%; Score 837; DB 16; Length 411;
 Best Local Similarity 100.0%; Pred. NO. 4.4e-54;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFRG 60
 DB 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFRG 60
 QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
 DB 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
 QY 121 PLVQECMVHDCADGKXPSPPEE 143
 DB 121 PLVQECMVHDCADGKXPSPPEE 143

RESULT 19
 AAR62999
 ID AAR62999 standard; protein; 411 AA.

XX AAR62999;
 XX 25-MAR-2003 (updated)
 DT 21-SEP-1995 (first entry)
 XX Pro-urokinase mutant His313.
 XX Pro-urokinase; thrombolysis; fibrin clot lysis; mutant His313;
 KW reduced fibrinogenolysis; non-specific plasminogen activation;
 KW systemic bleeding.
 XX Homo sapiens.

Key Location/Qualifiers
 Domain 297..313
 Disulfide-bond /note= "flexible loop"
 FT 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..279
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380

XX WO9511427-A1.
 XX 12-JAN-1995.
 XX 28-JUN-1994; 94WO-US07278.

PR 02-JUL-1993; 93US-0087163.
 XX (NEWB-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Gurewich V, Liu J;
 XX WPI; 1995-060991/08.
 XX Pro-urokinase mutants - have thrombolytic activity but reduced
 PT fibrinogenolysis activity and non-specific plasminogen activation
 XX Claim 11; Fig 1; 46pp; English.
 XX AAR62991 is the wild type pro-urokinase, from which the new mutants
 CC described in AAR62992-R63008 were derived. These mutants retain the
 CC thrombolytic activity of the wild type protein, useful for the
 CC treatment of thromboembolism, but have a reduced fibrinogenolysis
 CC activity and non-specific plasminogen activation. The mutants can
 CC therefore be used for the lysis of fibrin clots without inducing
 CC systemic bleeding, as can be the case with the wild type protein.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 411 AA;

Query Match 100.0%; Score 837; DB 16; Length 411;
 Best Local Similarity 100.0%; Pred. NO. 4.4e-54;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFRG 60
 DB 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFRG 60
 QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
 DB 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
 QY 121 PLVQECMVHDCADGKXPSPPEE 143
 DB 121 PLVQECMVHDCADGKXPSPPEE 143

RESULT 20
 AAR63000
 ID AAR63000 standard; protein; 411 AA.
 XX AAR63000;
 XX 25-MAR-2003 (updated)
 DT 21-SEP-1995 (first entry)
 XX Pro-urokinase mutant Ser175 His187.
 XX Pro-urokinase; thrombolysis; fibrin clot lysis;
 KW reduced fibrinogenolysis; non-specific plasminogen activation;
 KW systemic bleeding; mutant Ser175 His187.

Homo sapiens.
 Key Location/Qualifiers
 Domain 297..313
 Disulfide-bond /note= "flexible loop"
 FT 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..279
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380

```

XX  W09501427-A1.
PN
XX  Disulfide-bond 102..126
XX  Disulfide-bond 148..279
PD  12-JAN-1995.
XX
XX  Disulfide-bond 189..205
XX  Disulfide-bond 197..268
PF  28-JUN-1994; 94WO-US07278.
XX
XX  Disulfide-bond 293..362
XX  Disulfide-bond 325..341
XX  Disulfide-bond 352..380
PR  02-JUL-1993; 93US-0087163.
XX
XX  W09501427-A1.
PA  (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
XX
XX  Gurewich V, Liu J;
PI
XX
XX  WPI; 1995-060991/08.
DR
XX
XX  Pro-urokinase mutants - have thrombolytic activity but reduced
PT  fibrinogenolysis activity and non-specific plasminogen activation
XX
XX  Claim 15; Fig 1; 46pp; English.
PS
XX
XX  AAR62991 is the wild type pro-urokinase, from which the new mutants
CC  described in AAR62992-R63008 were derived. These mutants retain the
CC  thrombolytic activity of the wild type protein, useful for the
CC  treatment of thromboembolism, but have a reduced fibrinogenolysis
CC  activity and non-specific plasminogen activation. The mutants can
CC  therefore be used for the lysis of fibrin clots without inducing
CC  systemic bleeding, as can be the case with the wild type protein.
CC  (Updated on 25-MAR-2003 to correct PN field.)
XX
XX  Sequence 411 AA;
SQ
Query Match 100.0%; Score 837; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSDCCLNGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60
DB 1 SNELHQPNSDCCLNGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60
QY 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120
DB 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 PLVQECMVHDCADGKKPSPPEE 143
RESULT 21
AAR63001
ID AAR63001 standard; protein; 411 AA.
XX
XX  AAR63001;
AC
XX
XX  25-MAR-2003 (updated)
DT
XX  21-SEP-1995 (first entry)
DT
XX
XX  Pro-urokinase mutant Ser175 His187 Ala313.
DE
XX
XX  Pro-urokinase; thrombolysis; fibrin clot lysis;
KW  reduced fibrinogenolysis; non-specific plasminogen activation;
KW  systemic bleeding; mutant Ser175 His187 Ala313.
XX
XX  Homo sapiens.
OS
XX
XX  Key Location/Qualifiers
FH Domain 297..313
/note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113

```

```

FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX
XX  W09501427-A1.
PN
XX
XX  12-JAN-1995.
PD
XX
XX  28-JUN-1994; 94WO-US07278.
PF
XX
XX  02-JUL-1993; 93US-0087163.
PR
XX
XX  (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
PA
XX
XX  Gurewich V, Liu J;
PI
XX
XX  WPI; 1995-060991/08.
DR
XX
XX  Pro-urokinase mutants - have thrombolytic activity but reduced
PT  fibrinogenolysis activity and non-specific plasminogen activation
XX
XX  Claim 15; Fig 1; 46pp; English.
PS
XX
XX  AAR62991 is the wild type pro-urokinase, from which the new mutants
CC  described in AAR62992-R63008 were derived. These mutants retain the
CC  thrombolytic activity of the wild type protein, useful for the
CC  treatment of thromboembolism, but have a reduced fibrinogenolysis
CC  activity and non-specific plasminogen activation. The mutants can
CC  therefore be used for the lysis of fibrin clots without inducing
CC  systemic bleeding, as can be the case with the wild type protein.
CC  (Updated on 25-MAR-2003 to correct PN field.)
XX
XX  Sequence 411 AA;
SQ
Query Match 100.0%; Score 837; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSDCCLNGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60
DB 1 SNELHQPNSDCCLNGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHYRG 60
QY 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120
DB 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 PLVQECMVHDCADGKKPSPPEE 143
RESULT 22
AAR63002
ID AAR63002 standard; protein; 411 AA.
XX
XX  AAR63002;
AC
XX
XX  25-MAR-2003 (updated)
DT
XX  21-SEP-1995 (first entry)
DT
XX
XX  Pro-urokinase mutant Ser175 His187 Gly306.
DE
XX
XX  Pro-urokinase; thrombolysis; fibrin clot lysis;
KW  reduced fibrinogenolysis; non-specific plasminogen activation;
KW  systemic bleeding; mutant Ser175 His187 Gly306.
XX
XX  Homo sapiens.
OS
XX
XX  Key Location/Qualifiers
FH

```



```

AC AAR63004;
XX
XX
DT 25-MAR-2003 (updated)
DT 21-SEP-1995 (first entry)
XX
DE Pro-urokinase mutant Ser175 His187 Gly306 Ala313.
XX
XX Pro-urokinase; thrombolysis; fibrin clot lysis;
KW reduced fibrinogenolysis; non-specific plasminogen activation;
KW systemic bleeding; mutant Ser175 His187 Gly306 Ala313.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Domain 297..313
FT Disulfide-bond 11..19 /note= "flexible loop"
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX
XX WO9501427-A1.
XX
XX 12-JAN-1995.
XX
XX 28-JUN-1994; 94WO-US07278.
XX
XX 02-JUL-1993; 93US-0087163.
XX
XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
XX
XX Gurewich V, Liu J;
XX
XX WPI; 1995-060991/08.
XX
XX Pro-urokinase mutants - have thrombolytic activity but reduced
XX fibrinogenolysis activity and non-specific plasminogen activation
XX
XX Claim 15; Fig 1; 46pp; English.
XX
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
XX described in AAR62992-R63008 were derived. These mutants retain the
XX thrombolytic activity of the wild type protein, useful for the
XX treatment of thromboembolism, but have a reduced fibrinogenolysis
XX activity and non-specific plasminogen activation. The mutants can
XX therefore be used for the lysis of fibrin clots without inducing
XX systemic bleeding, as can be the case with the wild type protein.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 411 AA;
XX
XX Query Match 100.0%; Score 837; DB 16; Length 411;
XX Best Local Similarity 100.0%; Pred.No. 4.4e-54;
XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
XX |||||
XX 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
XX
XX 61 KASDTDMGRCLPWSATVLOQTYHAHRSALQGLGKHCNPNRRPWCYVQVGLK 120
XX |||||
XX 61 KASDTDMGRCLPWSATVLOQTYHAHRSALQGLGKHCNPNRRPWCYVQVGLK 120
XX
XX 121 PLVQECMVHDCADGKKPSPPEE 143
XX |||||

```

Db 121 PLVQECMVHDCADGKKPSPPEE 143

RESULT 25

AAR63005

ID AAR63005 standard; protein; 411 AA.

XX AAR63005;

XX 25-MAR-2003 (updated)

DT 21-SEP-1995 (first entry)

XX

DE Pro-urokinase mutant Ser175 His187 Ala300 Ala301 Ala313.

XX

KW Pro-urokinase; thrombolysis; fibrin clot lysis;
reduced fibrinogenolysis; non-specific plasminogen activation;
systemic bleeding; mutant Ser175 His187 Ala300 Ala313.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FH Domain 297..313

FT Disulfide-bond 11..19 /note= "flexible loop"

FT Disulfide-bond 13..31

FT Disulfide-bond 33..42

FT Disulfide-bond 33..42

FT Disulfide-bond 50..131

FT Disulfide-bond 71..113

FT Disulfide-bond 102..126

FT Disulfide-bond 148..279

FT Disulfide-bond 189..205

FT Disulfide-bond 197..268

FT Disulfide-bond 293..362

FT Disulfide-bond 325..341

FT Disulfide-bond 352..380

XX

XX WO9501427-A1.

XX

XX 12-JAN-1995.

XX

XX 28-JUN-1994; 94WO-US07278.

XX

XX 02-JUL-1993; 93US-0087163.

XX

XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.

XX

XX Gurewich V, Liu J;

XX

XX WPI; 1995-060991/08.

XX

XX Pro-urokinase mutants - have thrombolytic activity but reduced
fibrinogenolysis activity and non-specific plasminogen activation

XX

XX Claim 16; Fig 1; 46pp; English.

XX

XX AAR62991 is the wild type pro-urokinase, from which the new mutants
described in AAR62992-R63008 were derived. These mutants retain the
thrombolytic activity of the wild type protein, useful for the
treatment of thromboembolism, but have a reduced fibrinogenolysis
activity and non-specific plasminogen activation. The mutants can
therefore be used for the lysis of fibrin clots without inducing
systemic bleeding, as can be the case with the wild type protein.
(Updated on 25-MAR-2003 to correct PN field.)

XX

XX Sequence 411 AA;

XX

XX Query Match 100.0%; Score 837; DB 16; Length 411;

XX

XX Best Local Similarity 100.0%; Pred.No. 4.4e-54;

XX

XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

XX 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60

XX

XX 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60

XX

XX

CC activity and non-specific plasminogen activation. The mutants can
CC therefore be used for the lysis of fibrin clots without inducing
CC systemic bleeding, as can be the case with the wild type protein.
XX (Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 411 AA;
Query Match 100.0%; Score 837; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVOVGLK 120
DB 61 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVOVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 PLVQECMVHDCADGKKPSPPEE 143

RESULT 28

AAR63008
ID AAR63008 standard; protein; 411 AA.

AC AAR63008;

DT 25-MAR-2003 (updated)
DT 21-SEP-1995 (first entry)

DE Pro-urokinase mutant Ser175 His187 His300 Ala301 His313.

KW Pro-urokinase; thrombolysis; fibrin clot lysis;
KW reduced fibrinogenolysis; non-specific plasminogen activation;
KW systemic bleeding; mutant Ser175 His187 His300 Ala301 His313.

OS Homo sapiens.

Key Location/Qualifiers
Domain 297..313
FT /note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380

XX WO9501427-A1.

PN 12-JAN-1995.

XX 28-JUN-1994; 94WO-US07278.

XX 02-JUL-1993; 93US-0087163.

XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.

XX Gurewich V, Liu J;

XX WPI; 1995-060991/08.

XX Pro-urokinase mutants - have thrombolytic activity but reduced
XX fibrinogenolysis activity and non-specific plasminogen activation

XX Claim 16; Fig 1; 46pp; English.
PS

CC AAR62991 is the wild type pro-urokinase, from which the new mutants
CC described in AAR62992-R63008 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the
CC treatment of thromboembolism, but have a reduced fibrinogenolysis
CC activity and non-specific plasminogen activation. The mutants can
CC therefore be used for the lysis of fibrin clots without inducing
CC systemic bleeding, as can be the case with the wild type protein.
CC (Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 411 AA;

Query Match 100.0%; Score 837; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVOVGLK 120
DB 61 KASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVOVGLK 120
QY 121 PLVQECMVHDCADGKKPSPPEE 143
DB 121 PLVQECMVHDCADGKKPSPPEE 143

RESULT 29

AAR92926

ID AAR92926 standard; Protein; 411 AA.

XX AAR92926;

DT 03-AUG-1996 (first entry)

XX Pro-urokinase.

XX Pro-urokinase; plasminogen activator; fusion drug; drug delivery;
KW platelet; cardiovascular disease; thrombolytic.

OS Homo sapiens.

Key Location/Qualifiers
Region 1..132
FT /label= A-chain
FT Domain 1..45
FT Domain 46..132
FT Region 133..158
FT /label= Kringle_domain
FT /label= Linker_region
FT Cleavage-site 156..157
FT /note= "thrombin cleavage site"
FT Cleavage-site 158..159
FT /note= "plasmin cleavage site"
FT Region 159..411
FT /label= B-chain
FT Disulfide-bond 11
FT /note= "disulfide between Cys11 and Cys19"
FT Disulfide-bond 13
FT /note= "disulfide bond between Cys13 and Cys31"
FT Disulfide-bond 33
FT /note= "disulfide bond between Cys33 and Cys42"
FT Disulfide-bond 50
FT /note= "disulfide bond between Cys50 and Cys131"
FT Disulfide-bond 71
FT /note= "disulfide bond between Cys71 and Cys113"
FT Disulfide-bond 102
FT /note= "disulfide bond between Cys102 and Cys126"

FT Disulfide-bond 148 /note= "disulfide bond between Cys148 and Cys279"
 FT Disulfide-bond 189 /note= "disulfide bond between Cys189 and Cys205"
 FT Disulfide-bond 197 /note= "disulfide bond between Cys197 and Cys268"
 FT Disulfide-bond 293 /note= "disulfide bond between Cys293 and Cys362"
 FT Disulfide-bond 325 /note= "disulfide bond between Cys325 and Cys341"
 FT Disulfide-bond 352 /note= "disulfide bond between Cys352 and Cys380"
 XX WO9604004-A1.
 XX WO9604004-A1.
 XX 15-FEB-1996.
 XX 03-AUG-1995; 95WO-US09848.
 XX 05-AUG-1994; 94US-0286748.
 XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Gurewich V;
 XX WPI; 1996-129123/13.
 XX N-PSDB; AAT18237.
 XX Fusion product of plasminogen activator A chain and drug - targeted
 XX to platelets, useful for treatment of cardiovascular disease
 XX Claim 3; Page 39-40; 61pp; English.
 XX A portion of the A-chain (pref. amino acids 1-132) of pro-urokinase
 XX (AAR92926) can be linked to a drug, e.g. hirudin, somatostatin or
 XX their analogues (see also AAR92925 and AAR92927-33) and used in the
 XX prodn. of new fusion drugs. The constructs can be obtd. by expression
 XX of the appropriate nucleotide sequences in transformed host cells.
 XX When administered to a patient, the A-chain binds the fusion drug to
 XX the platelet outer membrane, i.e. to the site of thrombosis or
 XX vascular injury. Cleavage sites for thrombin and/or plasmin with
 XX the fusion drug allow the release of the drug at the target site.
 XX Sequence 411 AA;
 XX
 XX Query Match 100.0%; Score 837; DB 17; Length 411;
 XX Best Local Similarity 100.0%; Pred. No. 4.4e-54;
 XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
 XX 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
 XX
 XX 61 KASTDTNGRCLPWNSTVTLQOQTYHAHRSALQGLGKHNYCRPNRRPWCYVQVGLK 120
 XX 61 KASTDTNGRCLPWNSTVTLQOQTYHAHRSALQGLGKHNYCRPNRRPWCYVQVGLK 120
 XX
 XX 121 PLVQECMVHDCADGKFPSSPPEE 143
 XX 121 PLVQECMVHDCADGKFPSSPPEE 143
 XX
 XX RESULT 30
 XX JAY92836
 XX D AAY92836 standard; Protein; 411 AA.
 XX X
 XX C AAY92836;
 XX X
 XX T 29-AUG-2000 (first entry)
 XX X Urokinase plasminogen activator (uPA).
 XX E N-terminal; pro-uPA; urokinase plasminogen activator; receptor; uPAR;
 XX W

KW anti-cancer; anti-metastatic; anti-proliferative; anti-atherosclerotic;
 KW anti-thrombotic; anti-angiogenic; anti-inflammatory; anti-arthritis;
 KW anti-fibrotic; apoptotic; vasotropic; anti-diabetic; ophthalmological;
 XX thrombolytic.
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Disulfide-bond 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..279
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380
 XX WO200026353-A1.
 XX
 XX 11-MAY-2000.
 XX
 XX 28-OCT-1999; 99WO-US25210.
 XX
 XX 29-OCT-1998; 98US-0181816.
 XX
 XX (ANGS-) ANGSTROM PHARM INC.
 XX
 XX Mazar AP, Jones TR;
 XX
 XX WPI; 2000-365605/31.
 XX
 XX New cyclic peptide, useful for treatment or diagnosis of e.g. tumors
 XX and other diseases involving cell proliferation or migration, targets
 XX the urokinase plasminogen activator receptor
 XX
 XX Disclosure; Fig 1; 93pp; English.
 XX
 XX The present sequence shows the wild-type urokinase plasminogen activator
 XX (uPA). Cyclic peptides based on the amino acids residues 20-30 (the
 XX receptor-binding region) of uPA are claimed. These cyclic peptides target
 XX the uPA receptor (uPAR), allowing therapeutic or diagnostic agents to be
 XX delivered to uPAR-expressing cells. The cyclic peptides are used,
 XX optionally when linked to a therapeutic agent, to inhibit migration,
 XX invasion and proliferation of cells, or angiogenesis, or to induce
 XX apoptosis. Particularly they are used in human or veterinary medicine,
 XX to treat diseases characterized by these processes, e.g. solid tumors,
 XX leukaemia or lymphoma (or their metastases); benign hyperplasia;
 XX atherosclerosis; restenosis; ischaemia; deep vein thrombosis; neovascular
 XX glaucoma; diabetic retinopathy; arthritis; fibrosis; bone fracture etc.,
 XX most particularly growth, invasion and metastasis of tumors. When
 XX labeled, the cyclic peptides can be used for diagnostic detection of uPAR
 XX (a marker of metastasis) on cells, tissues etc., in vivo or in vitro, and
 XX when immobilized they are used to isolate uPAR or cells that express
 XX them. The cyclic peptides are stable, soluble in water, bind strongly to
 XX uPAR, are relatively inexpensive to produce and may be derivatized by
 XX attachment of therapeutic or diagnostic agents without significantly
 XX affecting their binding. Since they target uPAR, they should have
 XX relatively low systemic toxicity and only low doses are required.
 XX
 XX Sequence 411 AA;
 XX
 XX Query Match 100.0%; Score 837; DB 21; Length 411;
 XX Best Local Similarity 100.0%; Pred. No. 4.4e-54;
 XX Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
 XX 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60

Qy 61 KASTDTWGRPCLPWSATVLOQTYHAHRSDALQGLGKHNYCRNPDNRREPWCYVQVGLK 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 KASTDTWGRPCLPWSATVLOQTYHAHRSDALQGLGKHNYCRNPDNRREPWCYVQVGLK 120
Qy 121 PLVQECMVHDCADGKPSPEE 143
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 PLVQECMVHDCADGKPSPEE 143

Search completed: December 3, 2003, 14:39:15
Job time : 23.5259 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:35:43 ; Search time 4.96552 Seconds
(without alignments)
818.010 Million cell updates/sec

Title: US-09-880-503-9

Perfect score: 554
Sequence: 1 KTCYEGNGHFRGKASTDTM.....QECMVHDCADGKPKSSPPEE 96

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pdp.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pdp.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pdp.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pdp.*
5: /cgn2_6/ptodata/1/iaa/PTCUS.COMB.pdp.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	554	100.0	200	4	US-09-101-272G-73
2	554	100.0	208	4	US-09-101-272G-98
3	554	100.0	365	1	US-08-093-741-83
4	554	100.0	365	1	US-08-720-012-83
5	554	100.0	393	2	US-08-560-098A-44
6	554	100.0	393	3	US-08-967-024C-24
7	554	100.0	393	3	US-08-967-024C-25
8	554	100.0	411	1	US-08-087-163-1
9	554	100.0	411	1	US-08-286-748B-18
10	554	100.0	411	1	US-08-153-799-18
11	554	100.0	430	1	US-07-942-157A-3
12	554	100.0	431	4	US-09-101-272G-1
13	554	100.0	431	6	5188829-1
14	554	100.0	432	2	US-08-560-098A-47
15	544	98.2	411	3	US-09-181-816-1
16	543	98.0	411	2	US-08-560-098A-48
17	538.5	97.2	430	6	521969-2
18	530	95.7	157	3	US-08-142-590B-25
19	510	92.1	138	2	US-08-797-689-12
20	505	91.2	194	4	US-09-101-272G-80
21	505	91.2	201	4	US-09-101-272G-96
22	489	88.3	89	4	US-09-101-272G-62
23	241	43.5	477	2	US-08-560-098A-51
24	226	40.8	527	1	US-07-603-510B-16
25	226	40.8	527	2	US-08-811-949-39
26	226	40.8	527	5	PCT-US91-01025A-2
27	226	40.8	527	6	5185259-8

28	226	40.8	527	6	5520913-1	Patent No. 5520913
29	226	40.8	546	6	5200340-6	Patent No. 5200340
30	226	40.8	562	2	US-08-811-949-43	Sequence 43, Appl
31	226	40.8	562	2	US-08-560-098A-50	Sequence 50, Appl
32	226	40.8	562	2	US-08-883-795A-38	Sequence 38, Appl
33	226	40.8	562	6	5185259-3	Patent No. 5185259
34	226	40.8	562	6	5200340-2	Patent No. 5200340
35	226	40.8	562	6	5344773-2	Patent No. 5344773
36	221.5	40.0	356	1	US-08-427-640-8	Sequence 8, Appl
37	221	39.9	437	2	US-08-811-949-49	Sequence 49, Appl
38	221	39.9	437	2	US-08-811-949-51	Sequence 51, Appl
39	221	39.9	437	2	US-08-811-949-55	Sequence 55, Appl
40	221	39.9	437	2	US-08-811-949-57	Sequence 57, Appl
41	220.5	39.8	378	4	US-09-553-498-10	Sequence 10, Appl
42	220.5	39.8	378	4	US-09-618-869-10	Sequence 10, Appl
43	219.5	39.6	389	2	US-08-811-949-67	Sequence 67, Appl
44	218.5	39.4	326	4	US-09-411-977-3	Sequence 3, Appl
45	218.5	39.4	355	1	US-08-137-116-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-101-272G-73
; Sequence 73, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: 050979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: ATF domain of uPA
US-09-101-272G-73

Query Match	100.0%	Score 554;	DB 4;	Length 200;
Best Local Similarity	100.0%;	Pred. No. 2e-58;		
Matches	96;	Conservative 0;	Mismatches 0;	Indels 0;
Gaps	0;			
Qy	1	KTCYEGNGHFRGKASTDTMGRCLPWSATVLQOTYHAHRSALQGLGKHNCRPND	60	
Db	68	KTCYEGNGHFRGKASTDTMGRCLPWSATVLQOTYHAHRSALQGLGKHNCRPND	127	
Qy	61	RRRPWCYVQVGLKPLVQECMVHDCADGKPKSSPPEE	96	
Db	128	RRRPWCYVQVGLKPLVQECMVHDCADGKPKSSPPEE	163	

RESULT 2
US-09-101-272G-98
; Sequence 98, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: 050979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1

APPLICANT: WENNDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having Fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-44

Query Match 100.0%; Score 554; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 4.6e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60
|||||
Db 3 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 62
|||||
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96
|||||
Db 63 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 98
|||||

RESULT 6
US-08-967-024C-24
Sequence 24, Application US/08967024C
Patent No. 6133011
GENERAL INFORMATION:
APPLICANT: WENNDT, Stephan
APPLICANT: STEFFENS, Gerd Josef
APPLICANT: JANOSHA, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE: 30-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-24

Query Match 100.0%; Score 554; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 4.6e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60
|||||
Db 3 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 62
|||||
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96
|||||
Db 63 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 98
|||||

RESULT 7
US-08-967-024C-25
Sequence 25, Application US/08967024C
Patent No. 6133011
GENERAL INFORMATION:
APPLICANT: WENNDT, Stephan
APPLICANT: STEFFENS, Gerd Josef
APPLICANT: JANOSHA, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE: 30-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800

```

; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 100.0%; Score 554; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 4.8e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQO TYHAHRSDALQLGLGKHNYCRNPDN 60
Db 3 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQO TYHAHRSDALQLGLGKHNYCRNPDN 62

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
Db 63 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 98

RESULT 8
US-08-087-163-1
; Sequence 1, Application US/08087163
; Patent No. 5472692
; GENERAL INFORMATION:
; APPLICANT: Liu, Jian-Ning
; APPLICANT: Gurewich, Victor
; TITLE OF INVENTION: PRO-UKINASE MUTANTS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,163
; FILING DATE: 07/02/93
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04353/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
US-08-087-163-1

Query Match 100.0%; Score 554; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.8e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQO TYHAHRSDALQLGLGKHNYCRNPDN 60
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-967-024C-25

Query Match 100.0%; Score 554; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 4.8e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQO TYHAHRSDALQLGLGKHNYCRNPDN 60
Db 3 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQO TYHAHRSDALQLGLGKHNYCRNPDN 62

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
Db 63 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 98

RESULT 9
US-08-286-748B-18
; Sequence 18, Application US/08286748B
; Patent No. 5758542
; GENERAL INFORMATION:
; APPLICANT: Victor Gurewich
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
; OF DRUGS BY PLATELETS FOR THE TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,748B
; FILING DATE: August 5, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: J. Peter Fasse
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04547/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-286-748B-18

Query Match 100.0%; Score 554; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 4.8e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQO TYHAHRSDALQLGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQO TYHAHRSDALQLGLGKHNYCRNPDN 107

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 143

RESULT 10
US-08-153-799-18
; Sequence 18, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
```

```
;
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-153-799-18
;
; Query Match 100.0%; Score 554; DB 1; Length 411;
; Best Local Similarity 100.0%; Pred. No. 4.8e-58;
; Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLQOTYHAHRS DALQLGLGKHNCRNPDN 60
;
; Db 48 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLQOTYHAHRS DALQLGLGKHNCRNPDN 107
;
; QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPSPPEE 96
;
; Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPSPPEE 143
;
; RESULT 11
; US-07-942-157A-3
; Sequence 3, Application US/07942157A
; Patent No. 5648253
; GENERAL INFORMATION:
; APPLICANT: Wei, Cha-Mer
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
```

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,157A
; FILING DATE: 19920908
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631673
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: TSI108Cont.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..19
; OTHER INFORMATION: /label= peptide
; OTHER INFORMATION: /note= "WAP signal"
;
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 198..203
; OTHER INFORMATION: /label= modified
; OTHER INFORMATION: /note= "six amino acids deleted in mutant"
;
; US-07-942-157A-3
;
; Query Match 100.0%; Score 554; DB 1; Length 430;
; Best Local Similarity 100.0%; Pred. No. 5.1e-58;
; Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLQOTYHAHRS DALQLGLGKHNCRNPDN 60
;
; Db 67 KTCYEGNGHYRGKASTDTMGRCLPWN SATVLQOTYHAHRS DALQLGLGKHNCRNPDN 126
;
; QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPSPPEE 96
;
; Db 127 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPSPPEE 162
;
; RESULT 12
; US-09-101-272G-1
; Sequence 1, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (21)..()
; OTHER INFORMATION:
```

```

; NAME/KEY: misc_feature
; LOCATION: (20)..()
; OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)
US-09-101-272G-1

Query Match      100.0%; Score 554; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.1e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 128 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 163

RESULT 13
5188829-1
; Patent No. 5188829
; APPLICANT: KOBAYASHI, YO-ICHI; OMORI, MUNEKI; YAMADA, CHIKAKO
; TITLE OF INVENTION: RAPIDLY ACTING PROUROKINASE
; NUMBER OF SEQUENCES: 23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/340,007
; FILING DATE: 18-AUG-1988
; SEQ ID NO: 1
; LENGTH: 431
5188829-1

Query Match      100.0%; Score 554; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.1e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 128 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 163

RESULT 14
US-08-560-098A-47
; Sequence 47, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEZDI, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Everson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-47

Query Match      100.0%; Score 554; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 5.1e-58;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 69 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 128
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 129 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 164

RESULT 15
US-09-181-816-1
; Sequence 1, Application US/09181816
; Patent No. 6277818
; GENERAL INFORMATION:
; APPLICANT: MAZAR, Andrew P.
; APPLICANT: JONES, Terence R.
; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
; TITLE OF INVENTION: PLASMINOGEN ACTIVATOR RECEPTOR
; FILE REFERENCE: 32904200300 SIDN 1-7
; CURRENT APPLICATION NUMBER: US/09/181,816
; CURRENT FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-181-816-1

Query Match      98.2%; Score 544; DB 3; Length 411;
Best Local Similarity 99.0%; Pred. No. 7.5e-57;
Matches 95; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 143

RESULT 16
US-08-560-098A-48
; Sequence 48, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEZDI, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60

```

```
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
/ STREET: 1200 G Street, N.W., Suite 700
/ CITY: Washington
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/560,098A
/ FILING DATE: 17-NOV-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: P 44 40 892.7
/ FILING DATE: 17-NOV-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: EVANS, Joseph D.
/ REGISTRATION NUMBER: 26,269
/ REFERENCE/DOCKET NUMBER: 148/42448
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 628-8800
/ TELEFAX: (202) 628-8844
/ INFORMATION FOR SEQ ID NO: 48:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 411 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-560-098A-48

Query Match 98.0%; Score 543; DB 2; Length 411;
Best Local Similarity 99.0%; Pred. No. 9.9e-57;
Matches 95; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNP 60
Db 48 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNP 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 143

RESULT 17
5219569-2
; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO: 2:
; LENGTH: 430
5219569-2

Query Match 97.2%; Score 538.5; DB 6; Length 430;
Best Local Similarity 99.0%; Pred. No. 3.6e-56;
Matches 95; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNP 60
Db 68 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNP 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96

CORRESPONDENCE ADDRESS:
ADDRESSEE: Hibino, Tashihiro, Takahashi, Tadahito, Horii, Izumi; and
GOETINCK,
STREET: 128 RRRPWCYVQVGLKPLVQECMVHDCADG-KPSSPPEE 162
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,590B
FILING DATE: 25-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,318
FILING DATE: 02-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: MGP-009CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-142-590B-25

Query Match 95.7%; Score 530; DB 3; Length 157;
Best Local Similarity 95.8%; Pred. No. 1.1e-55;
Matches 92; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNP 60
Db 48 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNP 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 143

RESULT 18
US-08-142-590B-25
; Sequence 25, Application US/08142590B
; Patent No. 6120765
; GENERAL INFORMATION:
; APPLICANT: Hibino, Tashihiro, Takahashi, Tadahito, Horii, Izumi; and
GOETINCK,
; TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,590B
; FILING DATE: 25-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,318
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: MGP-009CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 157 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-142-590B-25

Query Match 95.7%; Score 530; DB 3; Length 157;
Best Local Similarity 95.8%; Pred. No. 1.1e-55;
Matches 92; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNP 60
Db 48 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNP 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 143

RESULT 19
US-08-797-689-12
; Sequence 12, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guillon, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
```

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Rhone-Poulenc Rorer Inc.
 STREET: 500 Arcola Road, 3C43
 CITY: Collegeville
 STATE: PA
 COUNTRY: USA
 ZIP: 19426
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Macintosh
 OPERATING SYSTEM: System 7.1
 SOFTWARE: Word 5.1 (Patent In)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/797,689
 FILING DATE: 31-JAN-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/256,927
 FILING DATE: 28-JUL-1994
 APPLICATION NUMBER: FR 92/01064
 FILING DATE: 31-JAN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/FR93/00085
 FILING DATE: 28-JAN-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Smith Ph.D., Julie K.
 REGISTRATION NUMBER: P-38,619
 REFERENCE/DOCKET NUMBER: ST92006-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610) 454-3839
 TELEFAX: (610) 454-3808
 INFORMATION FOR SEQ ID NO: 12:
 LENGTH: 138 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-797-689-12

Query Match 92.1%; Score 510; DB 2; Length 138;
 Best Local Similarity 100.0%; Pred. No. 2.3e-53;
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRPN 60
 DB 51 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRPN 110
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
 DB 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138

RESULT 20
 US-09-101-272G-80
 ; Sequence 80, Application US/09101272G
 ; Patent No. 6509445
 ; GENERAL INFORMATION:
 ; APPLICANT: Nissin Food Products Co., Ltd.
 ; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
 ; FILE REFERENCE: Q50979
 ; CURRENT APPLICATION NUMBER: US/09/101,272G
 ; PRIOR FILING DATE: 1998-07-08
 ; PRIOR APPLICATION NUMBER: JP 1059/1996
 ; NUMBER OF SEQ ID NOS: 107
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 80
 ; LENGTH: 194
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: ATPHI chimeric protein
 US-09-101-272G-80

Query Match 91.2%; Score 505; DB 4; Length 194;
 Best Local Similarity 100.0%; Pred. No. 1.4e-52;
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRPN 60
 DB 49 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRPN 108
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87
 DB 109 RRRPWCYVQVGLKPLVQECMVHDCADG 135

RESULT 21
 US-09-101-272G-96
 ; Sequence 96, Application US/09101272G
 ; Patent No. 6509445
 ; GENERAL INFORMATION:
 ; APPLICANT: Nissin Food Products Co., Ltd.
 ; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
 ; FILE REFERENCE: Q50979
 ; CURRENT APPLICATION NUMBER: US/09/101,272G
 ; PRIOR FILING DATE: 1998-07-08
 ; PRIOR APPLICATION NUMBER: JP 1059/1996
 ; NUMBER OF SEQ ID NOS: 107
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 96
 ; LENGTH: 201
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: ATPHI-CL chimeric protein
 US-09-101-272G-96

Query Match 91.2%; Score 505; DB 4; Length 201;
 Best Local Similarity 100.0%; Pred. No. 1.4e-52;
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRPN 60
 DB 49 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRPN 108
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87
 DB 109 RRRPWCYVQVGLKPLVQECMVHDCADG 135

RESULT 22
 US-09-101-272G-62
 ; Sequence 62, Application US/09101272G
 ; Patent No. 6509445
 ; GENERAL INFORMATION:
 ; APPLICANT: Nissin Food Products Co., Ltd.
 ; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
 ; FILE REFERENCE: Q50979
 ; CURRENT APPLICATION NUMBER: US/09/101,272G
 ; PRIOR FILING DATE: 1998-07-08
 ; PRIOR APPLICATION NUMBER: JP 1059/1996
 ; PRIOR FILING DATE: 1996-01-08
 ; NUMBER OF SEQ ID NOS: 107
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 62
 ; LENGTH: 89
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: residues 43-131 of the ATF domain of uPA
 US-09-101-272G-62

Query Match 88.3%; Score 489; DB 4; Length 89;

Best Local Similarity 100.0%; Pred. No. 4.4e-51;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNCRPN 60
DB 6 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNCRPN 65
QY 61 RRPWCYVQVGLKPLVQECMVHDC 84
DB 66 RRPWCYVQVGLKPLVQECMVHDC 89

RESULT 23
US-08-560-098A-51
; Sequence 51, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENDET, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-51

Query Match 43.5%; Score 241; DB 2; Length 477;
Best Local Similarity 50.0%; Pred. No. 1.3e-20;
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNCRPN 61
DB 127 TCYKQGVYRGWTSTSESGACINWNSLLRTYNGRSDAITLGLGNHNYCRPN 186
QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
DB 187 SRPWCYVIVKSKFILEFCVPCVCS 210

RESULT 24
US-07-609-510B-16
; Sequence 16, Application US/07609510B

; Patent No. 5326700
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,510B
; FILING DATE: 19901106
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-609-510B-16

Query Match 40.8%; Score 226; DB 1; Length 527;
Best Local Similarity 47.7%; Pred. No. 8.8e-19;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNCRPN 61
DB 91 TCYEDQGISYRGWTSTSESGAECTMNSALLAOKPYSGRRPDALRLGLGNHNYCRPN 150
QY 62 RRPWCYVQVGLKPLVQECMVHDC 87
DB 151 SKPWCYVFKAGKYSBFCSTPACSEG 176

RESULT 25
US-08-811-949-39
; Sequence 39, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OSLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-39

Query Match 40.8%; Score 226; DB 2; Length 527;
Best Local Similarity 47.7%; Pred. No. 8.8e-19;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 61
Db 91 TCYEDQGISYRGTWSTAESGAECTWNSSALAQKPYSGRRPDARLGLGNHNYCRNPDRD 150

QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
Db 151 SKPCYVFKAGKYSSEFCSTPACSEG 176

RESULT 26
PCT-US91-01025A-2
Sequence 2, Application PC/TUS9101025A
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Datin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/01025A
FILING DATE: 19910214
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/486,657
FILING DATE: 1 March 1990
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 454P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/266-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US91-01025A-2

Query Match 40.8%; Score 226; DB 5; Length 527;
Best Local Similarity 47.7%; Pred. No. 8.8e-19;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 61
Db 91 TCYEDQGISYRGTWSTAESGAECTWNSSALAQKPYSGRRPDARLGLGNHNYCRNPDRD 150

QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
Db 151 SKPCYVFKAGKYSSEFCSTPACSEG 176

RESULT 27
5185259-8
Patent No. 5185259
APPLICANT: GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
VEHAR, GORDON A.
TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
ACTIVATOR
NUMBER OF SEQUENCES: 15
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/489,855
FILING DATE: 02-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 12,694
FILING DATE: 09-FEB-1987
APPLICATION NUMBER: 483,052
FILING DATE: 07-APR-1983
APPLICATION NUMBER: 398,003
FILING DATE: 14-JUL-1982
APPLICATION NUMBER: 374,860
FILING DATE: 05-MAY-1982
SEQ ID NO: 8;
LENGTH: 527
5185259-8

Query Match 40.8%; Score 226; DB 6; Length 527;
Best Local Similarity 47.7%; Pred. No. 8.8e-19;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 61
Db 91 TCYEDQGISYRGTWSTAESGAECTWNSSALAQKPYSGRRPDARLGLGNHNYCRNPDRD 150

QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
Db 151 SKPCYVFKAGKYSSEFCSTPACSEG 176

RESULT 28
5520913-1
Patent No. 5520913
APPLICANT: ANDERSON, STEPHEN BENNETT, WILLIAM F.; BOTSTEIN,
DAVID; HIGGINS, DEBORAH L.; FAONI, NICHOLAS F.; ZOLLER, MARK J.
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR HAVING
ZMOGENIC PROPERTIES
NUMBER OF SEQUENCES: 35
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/88,451
FILING DATE: 06-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 770,510
FILING DATE: 03-OCT-1991
APPLICATION NUMBER: 384,608
FILING DATE: 24-JUL-1989
APPLICATION NUMBER: 240,856
FILING DATE: 02-SEP-1988
SEQ ID NO: 1;
LENGTH: 527
5520913-1

Query Match 40.8%; Score 226; DB 6; Length 527;
Best Local Similarity 47.7%; Pred. No. 8.8e-19;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:58 ; Search time 4.96552 Seconds
(without alignments)
1859.261 Million cell updates/sec

Title: US-09-880-503-9

Perfect score: 554

Sequence: 1 KTCYEGNGHFYRGKASTDTM.....QECMVHDCADGKKRSPPEE 96

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR 76: *
1: Pirl: *
2: Pirl2: *
3: Pirl3: *
4: Pirl4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Description
1	554	100.0	u-plasminogen acti
2	519	93.7	u-plasminogen acti
3	437.5	79.0	u-plasminogen acti
4	427	77.1	u-plasminogen acti
5	422	76.2	u-plasminogen acti
6	408	73.6	u-plasminogen acti
7	241	43.5	t-plasminogen acti
8	241	43.5	t-plasminogen acti
9	241	43.5	t-plasminogen acti
10	228.5	41.2	t-plasminogen acti
11	226	40.8	t-plasminogen acti
12	226	40.8	t-plasminogen acti
13	220	39.7	t-plasminogen acti
14	219	39.5	t-plasminogen acti
15	213	38.4	t-plasminogen acti
16	210	37.9	plasma hyaluronan-
17	199.5	36.0	plasma hyaluronan-
18	198.5	35.8	hepatocyte growth
19	194.5	35.1	coagulation factor
20	191.5	34.6	coagulation factor
21	170.5	30.8	apolipoprotein(a)
22	161	29.1	apolipoprotein(a)
23	157	28.3	plasmin (EC 3.4.21
24	156	28.2	plasmin (EC 3.4.21
25	155.5	28.1	coagulation factor
26	153	27.6	plasmin (EC 3.4.21
27	153	27.6	plasmin (EC 3.4.21
28	152	27.4	plasmin (EC 3.4.21
29	150	27.1	plasmin (EC 3.4.21

30 150 27.1 169 2 A40522 plasmin (EC 3.4.21
31 149 26.9 810 2 B30848 plasmin (EC 3.4.21
32 148.5 26.8 711 1 A47136 macrophage-stimula
33 148.5 26.8 810 2 I46260 plasmin (EC 3.4.21
34 148 26.7 790 1 PLPG plasmin (EC 3.4.21
35 147.5 26.6 716 1 A40332 macrophage-stimula
36 147.5 26.6 943 2 B45032 neurotrophic recep
37 146.5 26.4 810 1 PLHU plasmin (EC 3.4.21
38 146 26.4 937 2 A45082 neurotrophic recep
39 145 26.2 806 2 T18840 hypothetical prote
40 144.5 26.1 812 1 PLMS plasmin (EC 3.4.21
41 143 25.8 2869 2 T18518 apolipoprotein(a)
42 141.5 25.5 455 2 A61545 plasmin (EC 3.4.21
43 139 25.1 685 1 A48289 neurotrophic recep
44 137.5 24.8 716 1 JCS061 macrophage-stimula
45 137.5 24.8 728 1 A60185 hepatocyte growth

ALIGNMENTS

RESULT 1

UKHU

u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N:Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminoge
N:Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen ac
in form
C:Species: Homo sapiens (man)
C:Date: 17-Dec-1982 #sequence-revision 04-Dec-1986 #text-change 15-Sep-2000
C:Accession: A00931; I52209; J70102; A37561; I38102; S65783; A37562; A37563; A37564; A:
R:Riccio, A.; Grimaldi, G.; Verde, P.; Sebastiao, G.; Boast, S.; Blasi, F.
Nucleic Acids Res. 13, 2759-2771, 1985
A:Title: The human urokinase-plasminogen activator gene and its promoter.
A:Reference number: A00931; MUID:85215647; PMID:2987867
A:Accession: A00931
A:Molecule type: DNA
A:Residues: 1-431 <RIC>
A:Cross-references: GB:X02419; NID:G37601; PIDN:CAA26268.1; PID:G1834524
A:Note: the authors translated the codon ATG for residue 214 as Ile
R:Nagamine, Y.; Pearson, D.; Grattan, M.
Biochem. Biophys. Res. Commun. 132, 563-569, 1985
A:Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcine
A:Reference number: I52209; MUID:86050639; PMID:13933505
A:Accession: I52209
A:Status: Preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 145-161 <NAG1>
A:Cross-references: GB:X03027; NID:G340174; PIDN:AAA61257.1; PID:G340175
R:Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama,
Gene 36, 183-188, 1985
A:Title: Molecular cloning of cDNA coding for human preprourokinase.
A:Reference number: J70102; MUID:86056954; PMID:2415429
A:Accession: J70102
A:Molecule type: mRNA
A:Residues: 1-213, 'I', 215-431 <NAG2>
A:Cross-references: GB:X03226; NID:G340155; PIDN:AAC97138.1; PID:G340158; GB:D00244; N:
R:Verde, P.; Stoppe, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984
A:Title: Identification and primary sequence of an unspliced human urokinase poly(A) + I
A:Reference number: A37561; MUID:84272706; PMID:6589620
A:Accession: A37561
A:Molecule type: mRNA
A:Residues: 66-431 <VER>
A:Cross-references: GB:D00244; NID:G220138
R:Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elser
DNA 4, 139-146, 1985
A:Title: Molecular cloning, sequencing, and expression in Escherichia coli of human pre
A:Reference number: I38102; MUID:85203359; PMID:3888571
A:Accession: I38102
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>
A:Cross-references: EMBL:X02760; NID:G35297; PIDN:CAA26535.1; PID:G35298

R; Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, H. *Biochim. Biophys. Acta* 1293, 83-89, 1996

A; Title: Characterization of single chain urokinase-type plasminogen activator with a novel structure. *FEBS Lett.* 363, 1155-1165, 1992

A; Reference number: S65783; MUID: 96186279; PMID: 8652631

A; Accession: S65783

A; Status: preliminary

A; Molecule type: mRNA

A; Residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>

A; Cross-references: EMBL: D1143; NID: G1311467; PIDN: BAA01919.1; PID: g1199928

R; Gunzler, W.A.; Steffens, G.J.; Otting, F.; Kim, S.M.A.; Frankus, E.; Fliche, L. *Hoppe-Seyler's Z. Physiol. Chem.* 363, 1155-1165, 1992

A; Title: The primary structure of high molecular mass urokinase from human urine.

A; Reference number: A37562; MUID: 83055084; PMID: 6754569

A; Accession: A37562

A; Molecule type: protein

A; Residues: 21-177 <GUN>

R; Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O. *Eur. J. Biochem.* 125, 251-257, 1982

A; Title: Human low-molecular-weight urinary urokinase. Partial characterization and preliminary structure. *FEBS Lett.* 125, 251-257, 1982

A; Reference number: A37563; MUID: 8303608; PMID: 6749491

A; Accession: A37563

A; Molecule type: protein

A; Residues: 156-176, 179-193, 'T', 195, 'T', 197-224 <SCH>

R; Steffens, G.J.; Gunzler, W.A.; Otting, F.; Frankus, E.; Fliche, L. *Hoppe-Seyler's Z. Physiol. Chem.* 363, 1043-1058, 1992

A; Title: The complete amino acid sequence of low molecular mass urokinase from human urine.

A; Reference number: A37564; MUID: 83055099; PMID: 6754572

A; Accession: A37564

A; Molecule type: protein

A; Residues: 159-410 <STE>

R; Kentzer, E.J.; Buto, A.; Menon, G.; Sarin, V.K. *Biochem. Biophys. Res. Commun.* 171, 401-406, 1990

A; Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinant urokinase. *FEBS Lett.* 271, 1058-1064, 1990

A; Reference number: A35689; MUID: 90365737; PMID: 2393398

A; Accession: A35689

A; Molecule type: protein

A; Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>

A; Note: Identification of a fucose and attempt to determine its attachment site

R; Rabbani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazur, A.; Henkin, J.; Goltzman, B. *Biochim. Biophys. Res. Commun.* 173, 1058-1064, 1990

A; Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell line.

A; Reference number: A36697; MUID: 91097529; PMID: 2125213

A; Accession: A36697

A; Molecule type: protein

A; Residues: 21-34 <RAB>

R; Li, X.; Bokan, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M. *Submitted to the Brookhaven Protein Data Bank, July 1993*

A; Reference number: A51255; PDB: 1KDU

A; Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue

R; Li, X.; Smith, R.A.G.; Dobson, C.M. *Biochemistry* 31, 9562-9571, 1992

A; Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain of urokinase-type plasminogen activator. *J. Mol. Biol.* 222, 1-18, 1992

A; Reference number: A44375; MUID: 93003110; PMID: 1327118

A; Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR

R; Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettlesheim, D.G.; Mazar, A.P.; Olejniczak, Z. *Submitted to the Brookhaven Protein Data Bank, January 1994*

A; Reference number: A66822; PDB: 1URK

A; Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue

R; Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.; *Submitted to the Brookhaven Protein Data Bank, July 1995*

A; Reference number: A66058; PDB: 1LMW

A; Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426

C; Comment: This enzyme is found in urine in a high molecular mass form, consisting of A

C; Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a

C; Genetics:

A; Gene: GDB: PLAU

A; Cross-references: GDB: 119497; OMIM: 191840

A; Map position: 10q24-10q24

A; Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3

C; Function:

A; Description: proteolytically activates plasminogen

A; Pathway: fibrinolysis

C; Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C; Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine protease

F; 20/Domain: signal sequence #status predicted <SIG>

F; 21-431/Product: urokinase-type plasminogen activator, single chain form #status pred:

F; 21-177/Product: urokinase-type plasminogen activator chain A #status experimental <MI

F; 31-62/Domain: EGF homology <EGF>

F; 70-151/Domain: kringle homology <KRG>

F; 156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental

F; 179-433/Product: urokinase-type plasminogen activator chain B #status experimental <

F; 179-433/Domain: trypsin homology <TRY>

F; 31-39, 33-51, 53-62, 70-151, 91-133, 122-146, 168-299, 209-225, 217-288, 313-382, 345-361, 372-1

F; 38/Binding site: carboxylate (Thr) (covalent) #status predicted

F; 178-179/Cleavage site: Lys-Ile (plasmin) #status experimental

F; 224, 275, 376/Active site: His, Asp, Ser #status experimental

F; 322/Binding site: carboxylate (Asn) (covalent) #status experimental

Query Match 100.0%; Score 554; DB 1; Length 431;

Best Local Similarity 100.0%; Pred. No. 1.3e-51;

Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60

DB 58 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 96

DB 128 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 163

RESULT 2

URBA

u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon

C; Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)

C; Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999

C; Accession: S14687; S08651

R; Au, Y. P. T.; Wang, T. W.; Clowes, A. W. *Nucleic Acids Res.* 18, 3411, 1990

A; Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasmino

A; Reference number: S14687; MUID: 90287734; PMID: 2113276

A; Accession: S14687

A; Molecule type: mRNA

A; Residues: 1-433 <AU>

A; Cross-references: EMBL: X51935; NID: G38130; PIDN: CAA36200.1; PID: g38131

C; Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t

C; Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase

F; 21-176/Product: signal sequence #status predicted <SIG>

F; 30-61/Domain: EGF homology <EGF>

F; 69-150/Domain: kringle homology <KRG>

F; 178-433/Product: plasminogen activator chain B #status predicted <BCH>

F; 178-433/Domain: trypsin homology <TRY>

F; 167-298, 208-224, 216-287, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted

F; 223, 274, 378/Active site: His, Asp, Ser #status predicted

F; 324/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 93.7%; Score 519; DB 1; Length 433;

Best Local Similarity 94.8%; Pred. No. 7.4e-48;

Matches 91; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60

DB 67 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 126

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 96

DB 127 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 162

RESULT 3

URPG

u-plasminogen activator (EC 3.4.21.73) precursor - pig

C; Alternate names: uPA

C; Species: Sus scrofa domestica (domestic pig)

C; Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998

C:/Accession: A00932
R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
Nucleic Acids Res. 12, 9525-9541, 1984
A:Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A:Reference number: A00932; MUID:85087954; PMID:6096832
A:Accession: A00932
A:Molecule type: DNA
A:Residues: 1-240, 'H', 242-442 <NAG1>
A:Experimental source: kidney cell line LLC-PK1
R:Nagamine, Y.
submitted to the Protein Sequence Database, December 1986
A:Reference number: A37566
A:Contents: annotation; correction to residue 241
C:Genetics:
A:Introns: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
A:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:33-64/Domain: EGF homology <EGF>
F:72-153/Domain: kringle homology <KRG>
F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:190-430/Domain: trypsin homology <TRY>
F:152/Binding site: carboxylate (Asn) (covalent) #status predicted
F:179-310, 220-236, 228-299, 324-393, 356-372, 383-411/Disulfide bonds: #status predicted
F:235, 286, 387/Active site: His, Asp, Ser #status predicted

Query Match 79.0%; Score 437.5; DB 1; Length 442;
Best Local Similarity 74.3%; Pred. No. 3.9e-39;
Matches 78; Conservative 8; Mismatches 10; Indels 9; Gaps 1;
Qy 1 KTCYEGNGHFYRGKASTDWTGRCPLPWNSTVLTQTYHAHRSDALQGLGKHNYCRNPDN 60
Db 70 QTCFEGNGSHYRGKANTDGTGRPCLPWNSTVLTQTYHAHRSDALQGLGKHNYCRNPDN 129
Qy 61 RRRPWCYVQGLKPLVQECMVHDCADGCKKSPPEE 96
Db 130 QRRPWCYVQGLKPLVQECMVHDCADGCKKSPPEE 174

RESULT 4
S18932
u-plasminogen activator (EC 3.4.21.73) precursor - rat
N:Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator
C:Species: Rattus norvegicus (Norway rat)
C>Date: 18-Oct-1989 #sequence revision 10-Feb-1995 #text_change 18-Jun-1999
C:Accession: S24604; I50186; I53472; S18932
R:Kabbani, S.A.
submitted to the EMBL Data Library, April 1992
A:Reference number: S24604
A:Accession: S24604
A:Molecule type: mRNA
A:Residues: 1-15, 'H', 17-23, 'G', 25-331, 'N', 333-432 <RAB>
A:Cross-references: EMBL:X65851; NID:957456; PIDN:CAA46601.1; PID:957457
A:Experimental source: tissue kidney
R:Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.
Cancer Res. 52, 2489-2496, 1992
A:Title: Transcriptional and posttranscriptional activation of urokinase plasminogen act
A:Reference number: I60186; MUID:92233409; PMID:1568219
A:Accession: I60186
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-432 <RES>
A:Cross-references: EMBL:X63434; NID:957465; PIDN:CAA45028.1; PID:957466
A:Experimental source: strain Fischer 344; tissue mammary
R:Ragno, P.; Cassano, S.; Degen, J.; Kessler, C.; Blasi, F.; Rossi, G.
FEBS Lett. 306, 193-198, 1992
A:Title: The receptor for the plasminogen activator of urokinase type is up-regulated in
A:Reference number: I53472; MUID:92339349; PMID:1321734
A:Accession: I53472
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 31-62 <RE2>

A:Cross-references: EMBL:X66907; NID:G396200; PIDN:CAA47356.1; PID:G938279
C:Genetics:
A:Gene: uPA
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:31-62/Domain: EGF homology <EGF>
F:70-151/Domain: kringle homology <KRG>
F:179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:179-420/Domain: trypsin homology <TRY>
F:168-300, 210-226, 314-383, 346-362, 373-401/Disulfide bonds: #status predicted
F:225, 276, 377/Active site: His, Asp, Ser #status predicted

Query Match 77.1%; Score 427; DB 1; Length 432;
Best Local Similarity 77.1%; Pred. No. 5.1e-38;
Matches 74; Conservative 7; Mismatches 15; Indels 0; Gaps 0;
Qy 1 KTCYEGNGHFYRGKASTDWTGRCPLPWNSTVLTQTYHAHRSDALQGLGKHNYCRNPDN 60
Db 68 KTCYHGNGSQSYRGKANTDGTGRPCLPWNSTVLTQTYHAHRSDALQGLGKHNYCRNPDN 127
Qy 61 RRRPWCYVQGLKPLVQECMVHDCADGCKKSPPEE 96
Db 128 QRRPWCYVQGLKPLVQECMVHDCADGCKKSPPEE 163

RESULT 5
u-plasminogen activator (EC 3.4.21.73) precursor - bovine
N:Alternate names: uPA
C:Species: Bos primigenius taurus (cattle)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C:Accession: JN0560
R:Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
Gene 125, 177-183, 1993
A:Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and indi
A:Reference number: JN0560; MUID:93216113; PMID:18385052
A:Accession: JN0560
A:Molecule type: mRNA
A:Residues: 1-433 <KRA>
A:Cross-references: GB:L03546; NID:G163800; PIDN:AAA51419.1; PID:G163801
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-179/Product: plasminogen activator chain A #status predicted <MA1>
F:21-179/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:33-64/Domain: EGF homology <EGF>
F:72-153/Domain: kringle homology <KRG>
F:181-433/Product: plasminogen activator chain B #status predicted <MA2>
F:181-431/Domain: trypsin homology <TRY>
F:170-301, 211-227, 219-230, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
F:225, 277, 378/Active site: His, Asp, Ser #status predicted

Query Match 76.2%; Score 422; DB 1; Length 433;
Best Local Similarity 75.0%; Pred. No. 1.8e-37;
Matches 72; Conservative 9; Mismatches 15; Indels 0; Gaps 0;
Qy 1 KTCYEGNGHFYRGKASTDWTGRCPLPWNSTVLTQTYHAHRSDALQGLGKHNYCRNPDN 60
Db 70 KTCYQNGSHYRGKANDLSGRPCLPWNSTVLTQTYHAHRSDALQGLGKHNYCRNPDN 129
Qy 61 RRRPWCYVQGLKPLVQECMVHDCADGCKKSPPEE 96
Db 130 QRRPWCYVQGLKPLVQECMVHDCADGCKKSPPEE 165

RESULT 6
u-plasminogen activator (EC 3.4.21.73) precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
C:Accession: A29420; A24615

R;Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
 Biochemistry 26, 8270-8279, 1987
 A;Title: The murine urokinase-type plasminogen activator gene.
 A;Reference number: A29420; MUID:88163489; PMID:2831940
 A;Accession: A29420
 A;Molecule type: DNA
 A;Residues: 1-433 <DEG>
 A;Cross-references: GB:M17922; NID:9202296; PIDN:AAA40539.1; PID:9202297
 R;Beilin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher, Eur. J. Biochem. 148, 225-232, 1995
 A;Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase
 A;Reference number: A24615; MUID:85179474; PMID:2985383
 A;Accession: A24615
 A;Molecule type: mRNA
 A;Residues: 1-433 <BEL>
 A;Cross-references: GB:X02389; NID:955127; PIDN:CAA26231.1; PID:955129
 C;Genetics:
 A;Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
 C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F;1-20/Domain: signal sequence #status predicted <SIG>
 F;21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F;32-63/Domain: EGF homology <EGF>
 F;71-152/Domain: kringle homology <KRG>
 F;180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F;180-421/Domain: trypsin homology <TRY>
 F;169-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
 F;226,277,378/Active site: His, Asp, Ser #status predicted
 Query Match 73.6%; Score 408; DB 1; Length 433;
 Best Local Similarity 70.8%; Pred. No. 5.5e-36;
 Matches 69; Conservative 12; Mismatches 16; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQGLGKHNCRPNPN 60
 DB 69 KTCYHGNGDSYRGKANTDKRGRCPLAWNAPAVLQKYPNAHRPDAISLGLGKHNCRPNPN 128
 QY 61 RRPWCYVQVGLKPLVQECMVHDCADGKXPSPPEE 96
 DB 129 QKRPWCYVQIGLRQVQECMVHDCSLSKPSSVDQ 164
 RESULT 7
 JS0599
 t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
 N;Alternate names: tissue plasminogen activator
 C;Species: Desmodus rotundus (common vampire bat)
 C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C;Accession: JS0599
 R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Dor
 Gene 105, 229-237, 1991
 A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des
 A;Reference number: JS0597; MUID:92039036; PMID:1937019
 A;Accession: JS0599
 A;Molecule type: mRNA
 A;Residues: 1-431 <KRA>
 A;Cross-references: GB:M63989; NID:G166076; PIDN:AAA31594.1; PID:G166077
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-36/Domain: propeptide #status predicted <PRO>
 F;37-431/Product: plasminogen activator beta #status predicted <PLA>
 F;41-74/Domain: EGF homology <EGF>
 F;82-163/Domain: kringle homology <KRG>
 F;180-425/Domain: trypsin homology <TRY>
 F;41-52,45-63,65-74,82-163,103-135,134-158,168-299,211-227,219-288,313-388/Disulfide bon
 F;139,352/Binding site: carboxylate (Asn) (covalent) #status predicted
 F;179-180/Cleavage site: His-Ser (plasma) #status predicted
 F;226,275,382/Active site: His, Asp, Ser #status predicted
 F;345-361,378-406/Disulfide bonds: #status predicted
 Query Match 43.5%; Score 241; DB 2; Length 431;
 Best Local Similarity 50.0%; Pred. No. 4.1e-18;

Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;
 QY 2 TCYEGNGHYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQGLGKHNCRPNPN 61
 DB 127 TCYKDGQVYRGTWSTSSGACINMNSNLLTRTYNGRRSDAITLGLGNHNYCRPNPN 186
 QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
 DB 141 SKPWCYVIKASKFILEFCSPVCS 164
 RESULT 8
 JS0598
 t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
 C;Species: Megaderma lyra
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C;Accession: A34369
 R;Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jack
 J. Biol. Chem. 264, 17947-17952, 1989
 A;Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmi
 A;Reference number: A34369; MUID:90036867; PMID:2509450
 A;Accession: A34369
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-477 <GAR>
 A;Cross-references: GB:J05082; NID:G166080; PIDN:AAA31596.1; PID:G166081
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat ho
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-36/Domain: propeptide #status predicted <PRO>
 F;37-477/Product: plasminogen activator #status predicted <PLA>
 F;87-120/Domain: fibronectin type I repeat homology <1FA>
 F;128-209/Domain: EGF homology <EGF>
 F;226-471/Domain: trypsin homology <TRY>
 F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-
 F;272,321,428/Active site: His, Asp, Ser #status predicted
 Query Match 43.5%; Score 241; DB 1; Length 477;
 Best Local Similarity 50.0%; Pred. No. 4.5e-18;
 Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;
 QY 2 TCYEGNGHYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQGLGKHNCRPNPN 61
 DB 127 TCYKDGQVYRGTWSTSSGACINMNSNLLTRTYNGRRSDAITLGLGNHNYCRPNPN 186
 QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
 DB 187 SKPWCYVIKASKFILEFCSPVCS 210
 RESULT 9
 JS0598
 t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
 N;Alternate names: tissue plasminogen activator
 C;Species: Desmodus rotundus (common vampire bat)
 C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C;Accession: JS0598
 R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; D
 Gene 105, 229-237, 1991
 A;Title: The plasminogen activator family from the salivary gland of the vampire bat D
 A;Reference number: JS0597; MUID:92039036; PMID:1937019
 A;Accession: JS0598
 A;Molecule type: mRNA
 A;Residues: 1-477 <KRA>
 A;Cross-references: GB:M63988; NID:G166074; PIDN:AAA31593.1; PID:G166075
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat ho
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-36/Domain: propeptide #status predicted <PRO>
 F;37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>
 F;42-79/Domain: fibronectin type I repeat homology <1FA>
 F;87-120/Domain: EGF homology <EGF>

```
F:128-209/Domain: kringle homology <KRG>
F:226-471/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F:185,398/Binding site: carbohydrate (Asn) #status predicted
F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match          43.5%; Score 241; DB 2; Length 477;
Best Local Similarity 50.0%; Pred. No. 4.5e-18;
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTWGRCLPNSATVLCQTYHAHRSALQGLGKXNYCRNPDR 61
DQ 127 TCYKQGVYRGTTWTSSEGAICINWNSLLRTYNGRSDAITLGLGNHNYCRNPDR 186
QY 62 RRPWCYVQGLKPLVQECMVHCA 85
DQ 187 SRPWCYVIRKASPILEFCSPVCS 210

RESULT 10
A35029
t-plasminogen activator (EC 3.4.21.68) precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A35029; A31597
R:Feng, P.; Ohlsson, M.; NY, T.
J. Biol. Chem. 265, 2022-2027, 1990
A:Title: The structure of the TARA-less rat tissue-type plasminogen activator gene. Spec
A:Reference number: A35029; MUID:90130448; PMID:2105315
A:Accession: A35029
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-559 <FEN>
A:Cross-references: GB:M31197; NID:G207429; PIDN:AAA42261.1; PID:G207431; GB:J05226
R:Ny, T.; Leonardson, G.; Hsueh, A.J.W.
DNA 7, 671-677, 1988

A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activator
A:Reference number: A31597; MUID:89170114; PMID:3148445
A:Accession: A31597
A:Molecule type: mRNA
A:Residues: 1-379, 'K', 381-559 <NVT>
A:Cross-references: GB:M23697; NID:G530159; PIDN:AAA41812.1; PID:G530160
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-29/Domain: propeptide #status predicted <PRO>
F:30-359/Product: t-plasminogen activator #status predicted <MAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F:38-75/Domain: fibronectin type I repeat homology <IF1>
F:83-116/Domain: EGF homology <EGF>
F:124-205/Domain: kringle homology <KR1>
F:213-294/Domain: kringle homology <XR2>
F:309-559/Product: t-plasminogen activator <TRY>
F:309-553/Domain: trypsin homology <TRY>
F:38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4
F:149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match          41.2%; Score 228.5; DB 1; Length 559;
Best Local Similarity 45.3%; Pred. No. 1.1e-16;
Matches 43; Conservative 11; Mismatches 36; Indels 5; Gaps 1;

QY 2 TCYEGNGHYRGKASTDTWGRCLPNSATVLCQTYHAHRSALQGLGKXNYCRNPDR 61
DQ 123 TCYEGQGITVRGTTWSTAENGACINWNSALSOKPYARRPNAIKLGLGNHNYCRNPDR 182
QY 62 RRPWCYVQGLKPLVQECMVHDCADGKXSSPPEE 96
DQ 183 VKPWCYVFRKAGKYTFEFCSTPAC-----PKGPTED 212
```

```
RESULT 11
I38098
t-plasminogen activator precursor, inactive endothelial splice form - human
N:Alternate names: t-tissue plasminogen activator
C:Species: Homo sapiens (man)
C>Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
C:Accession: I38098; S01678
R:Siebert, P.D.; Fong, K.
Nucleic Acids Res. 18, 1086, 1990
A:Title: Variant tissue type plasminogen activator (PLAT) cDNA obtained from human endo
A:Reference number: I38098; MUID:90192128; PMID:1969145
A:Accession: I38098
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-291 <SIE>
A:Cross-references: EMBL:X13097; NID:G35282; PIDN:CAA1489.1; PID:G35283
C:Comment: For the main splice form, see PIR:UKHUT. This form probably does not have pr
C:Genetics:
A:Gene: GDB:PLAT
A:Cross-references: GDB:119496; OMIM:173370
A:Map position: 8p12-8p12
A:Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat ho
C:Keywords: alternative splicing; fibrinolysis; glycoprotein; kringle
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-32/Domain: propeptide #status predicted <PRO>
F:33-291/Product: t-plasminogen activator, inactive endothelial splice form #status pre
F:41-78/Domain: fibronectin type I repeat homology <FPA>
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology <KR1>
F:215-291/Domain: kringle homology #status atypical <KR2>
F:41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203/Disulfide bonds: #status pre

Query Match          40.8%; Score 226; DB 2; Length 291;
Best Local Similarity 47.7%; Pred. No. 1.1e-16;
Matches 41; Conservative 8; Mismatches 37; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTWGRCLPNSATVLCQTYHAHRSALQGLGKXNYCRNPDR 61
DQ 126 TCYEDQGITVRGTTWSTAENGACINWNSALSOKPYARRPNAIKLGLGNHNYCRNPDR 185
QY 62 RRPWCYVQGLKPLVQECMVHDCADG 87
DQ 186 SKPWCYVFRKAGKYSSEFCSTPACSEG 211

RESULT 12
UKHUT
t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
N:Alternate names: t-PA; tissue plasminogen activator
C:Species: Homo sapiens (man)
C>Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 08-Dec-2000
C:Accession: A94004; A23529; J70562; A93293; S02125; A91343; A93951; A91322; A54645; I
R:Ny, T.; Elgh, F.; Lund, B.
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
A:Title: The structure of the human tissue-type plasminogen activator gene: correlatio
A:Reference number: A94004; MUID:84298137; PMID:6089198
A:Accession: A94004
A:Molecule type: DNA
A:Residues: 1-562 <NYT>
A:Cross-references: GB:L00141
A:Note: the codon given for residue 93 (ACC) is inconsistent with the authors' transla
R:Friezner Degen, S.J.; Rajput, B.; Reich, E.
J. Biol. Chem. 261, 6972-6985, 1986
A:Title: The human tissue plasminogen activator gene.
A:Reference number: A23529; MUID:86196143; PMID:3009482
A:Accession: A23529
A:Molecule type: DNA
A:Residues: 1-562 <DEG>
A:Cross-references: GB:K03021; NID:G339817; PIDN:AAA98809.1; PID:G339818
R:Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuoka, S.; Higashio, K.
Agric. Biol. Chem. 55, 1225-1232, 1991
A:Title: Purification and characterization of tissue plasminogen activator secreted by
```

A:Reference number: JT0562; MUID:91291340; PMID:1368691
A:Accession: JT0562
A:Molecule type: mRNA
A:Residues: 31-562 <ITA>
A:Cross-references: DBJ:001086; NID:920128; PIDN:BAAC00881.1; PID:g441174
A:Experimental source: embryonic lung fibroblast IMR-90 cells
A:Note: part of this sequence, including the amino end of the mature protein, was confirmed by R. Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vohar, G.A.; Ward, C.A.; Bennett Nature 301, 214-221, 1983
A:Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Escherichia coli
A:Reference number: A93293; MUID:83115262; PMID:6337343
A:Accession: A93293
A:Molecule type: mRNA
A:Residues: 1-562 <PNS>
A:Cross-references: GB:L00141
A:Experimental source: melanoma cells
R:Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M. Nucleic Acids Res. 16, 5695, 1988
A:Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fetal lung
A:Reference number: S02125; MUID:88262579; PMID:3133640
A:Accession: S02125
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-562 <SAS>
A:Cross-references: EMBL:X07293; NID:g37243; PIDN:CAA30302.1; PID:g37244
A:Experimental source: fetal lung cells
R:Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Matsuda, T. FEBS Lett. 189, 145-149, 1985
A:Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen activator cDNA
A:Reference number: A91343; MUID:85285620; PMID:3896853
A:Accession: A91343
A:Molecule type: mRNA
A:Residues: 1-38, 'G', 'E', '435-562' <XAG>
A:Experimental source: Detroit 562 cells; ATCC 138
R:Edlund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S. Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
A:Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activator
A:Reference number: A93951; MUID:83169656; PMID:6572897
A:Accession: A93951
A:Molecule type: mRNA
A:Residues: 251-358 <EDL>
A:Experimental source: melanoma cells
R:Fehl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H. Biochemistry 23, 3701-3707, 1984
A:Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived amino acid sequence
A:Reference number: A90488; MUID:85000468; PMID:6433976
A:Contents: annotation; melanoma cells; partial sequence of residues 36-562, active and inactive
R:Fehl, G.; Kaplan, L.; Enarsson, M.; Wallen, P.; Jornvall, H. FEBS Lett. 168, 29-32, 1984
A:Title: Differences between uterine and melanoma forms of tissue plasminogen activator
A:Reference number: A91322; MUID:84158956; PMID:6538514
A:Accession: A91322
A:Molecule type: protein
A:Residues: 33-45; 311-320 <POH>
A:Experimental source: uterus
A:Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
R:Van Zonneveld, A.J.; Veerman, H.; Pannekoek, H. J. Biol. Chem. 261, 14214-14218, 1986
A:Reference number: A37567; MUID:87033611; PMID:3021732
A:Contents: annotation; fibrin binding site
R:Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Engelen, J. J. 3525-3530, 1986
A:Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen activator in fibrinolysis
A:Reference number: A37568; MUID:87161761; PMID:3030730
A:Contents: annotation; fibrin binding site
R:Dodd, I.; Nunn, B.; Robinson, J.H. Thromb. Haemost. 59, 523-528, 1988
A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type plasminogen activator
A:Reference number: A60902; MUID:89044681; PMID:3142086
A:Contents: annotation; novel forms of expressed recombinant t-PA
R:Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emrige, J.S.; Odenakker, G.; Mor. Biol. Med. 3, 279-292, 1986

A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its expression in Escherichia coli
A:Reference number: A54645; MUID:86284200; PMID:3050401
A:Accession: A54645
A:Molecule type: mRNA
A:Residues: 1-562 <HAR>
A:Cross-references: GB:M15518; NID:g190031; PIDN:AAA60111.1; PID:g190032
A:Note: parts of this sequence were confirmed by peptide sequencing
R:Reddy, V.B.; Garramane, A.J.; Sasak, H.; Wei, C. DNA 6, 461-472, 1987
A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells
A:Reference number: I60110; MUID:88054470; PMID:2824147
A:Accession: I60110
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-562 <RES>
A:Cross-references: GB:M18182; NID:g340176; PIDN:AAA36800.1; PID:g340177
R:Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D. J. Biol. Chem. 260, 11233-11230, 1985
A:Title: Isolation and characterization of the human tissue-type plasminogen activator cDNA
A:Reference number: I55232; MUID:85289338; PMID:3161693
A:Accession: I55232
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-36 <RE2>
A:Cross-references: GB:M11890; NID:g339837; PIDN:AAA61213.1; PID:g339839
C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a single disulfide bond
C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond
C:Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat
C:Genetics:
A:Gene: GDB:PLAT
A:Map position: 8p12-8p12
A:Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 500/1
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-32/Domain: propeptide #status predicted <PRO>
F:33-562/Product: t-plasminogen activator #status experimental <MAT>
F:33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
F:41-78/Domain: fibronectin type I repeat homology <IFI>
F:86-119/Domain: EGF homology <EGF>
F:127-208/Domain: kringle homology <KR1>
F:215-296/Domain: kringle homology <KR2>
F:311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
F:311-556/Domain: trypsin homology <TRY>
F:41-71-69-78-96-97-91-108-110-119-127-208-148-190-179-203-215-296-236-278-267-291-299-152-483/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:219/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F:310-311/Cleavage site: Arg-ile (plasmin, trypsin) #status experimental
F:357-406/Active site: His, Asp #status predicted
F:513/Active site: Ser #status experimental

Query Match 40.8%; Score 226; DB 1; Length 562;
Best Local Similarity 47.7%; Pred. No. 2, le-16;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYENGHFYFGKASTMTGRCPLPWSATVLQOQTYAHRSDALQGLGKHNYCRPDNR 61
DB 126 TCYEDQGISYRGTSWTAESGAECTNMWNSALQAQPKSYGRPRDAIRLGLGHNHNYCRPDNR 185
QY 62 RRPWCYQVQGLKPLVQECMVHDCADG 87
DB 186 SKPCWYVFKAGKYSERFSTPACSEG 211

RESULT 13
JS0600
t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0600
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; De

F;149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F;355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 39.5%; Score 219; DB 1; Length 559;
Best Local Similarity 46.0%; Pred. No. 1,2e-15;
Matches 40; Conservative 11; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRFGKASTDTMGRPLCPNNSATVLOOQTYHAHRSDALQLGLGKHNYCRNPDR 61
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
DB 123 TCFEEQGITVRGTWSTAESGAECINNSSLKPYNARRPNAIKLGLGNHNYCRNPDRD 182
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 62 RRPWCYVQVLKPLVQECMVHDCAQG 88
: ||| : ||| : ||| : ||| : ||| : ||| : ||| :
DB 183 LKPWCYVFRAKGYTTTFCSTFPACPKGK 209
: ||| : ||| : ||| : ||| : ||| : ||| : ||| :

RESULT 15
JS0597
N:Plasminogen activator (EC 3.4.21.66) alpha-1 precursor - common vampire bat
t_Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (Common vampire bat)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0597
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Dörmann, C.
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat
A:Reference number: JS0597; PMID:92039036; PMID:1937019
A:Accession: JS0597
A:Molecule type: mRNA
A:Residues: 1-477 <XRA>
C:Cross-references: GB:M63987; NID:G166070; PIDN:AAA31591.1; PID:G166071
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hct
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <lFA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: kringle homology <KRK>
F:226-471/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359,
F:453,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 38.4%; Score 213; DB 2; Length 477;
Best Local Similarity 46.4%; Pred. No. 4.4e-15;
Matches 39; Conservative 10; Mismatches 35; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRFGKASTDTMGRPLCPNNSATVLOOQTYHAHRSDALQLGLGKHNYCRNPDR 61
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
DB 127 TCVGGGVTVRGTWSTAESRVEICINNSSLTRRTYNGRMPEAFNLGLGNHNYCRNPNGA 186
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 62 RRPWCYVQVLKPLVQECMVHDCA 85
: ||| : ||| : ||| : ||| : ||| : ||| : ||| :
DB 187 PKPWCVIRAKGYTSCTSPVCVS 210
: ||| : ||| : ||| : ||| : ||| : ||| : ||| :

RESULT 16
JC5878
plasma hyaluronan-binding protein precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 16-Jul-1999
C:Accession: JC5878
R:Hashimoto, K.; Tohe, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura, H.
Biol. Pharm. Bull. 20, 1127-1130, 1997
A:Title: Cloning of the cDNA for a mouse homologue of human PHBP: A novel hyaluronan-bi
A:Reference number: JC5878; MUID:98065239; PMID:9401717
A:Accession: JC5878
A:Molecule type: mRNA
A:Residues: 1-558 <HAS>
C:Comment: This protein acts as serine protease.

C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; tryptophan
 F:1-231/Domain: signal sequence #status predicted <SIG>
 F:24-311/Product: plasma hyaluronan-binding protein large chain #status predicted <MATL>
 F:75-106/Domain: EGF homology <EG1>
 F:113-145/Domain: EGF homology <EG2>
 F:152-185/Domain: EGF homology <EG3>
 F:192-274/Domain: kringle homology <KRI>
 F:312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <MATS>
 F:312-548/Domain: trypsin homology <TRY>

Query Match 37.9%; Score 210; DB 2; Length 556;
 Best Local Similarity 43.5%; Pred. No. 1.4e-14;
 Matches 40; Conservative 14; Mismatches 34; Indels 4; Gaps 2;

QY 3 CYENGHFYRGKASTDTMGRCPLPWSNATVLTQYTHAHRSDALQLGLGKHNYCRNPDR 62

DB 192 CYVGDGYSYRGKSKVTNQPCLYWNSHLLQETYNFMEDAEATHGIAHFCRNPDPGH 251

QY 63 RPYCVVOGLKPLVQE-CWVHDCADGKKPSP 93

DB 252 KPWCFVKNSKWEYCDVTVC---FVPTDP 280

RESULT 17

JC4795
 plasma hyaluronan-binding protein precursor - human
 N:Alternate names: hepatocyte growth factor activator-like protein; PHBP
 N:Contains: serine proteinase (EC 3.4.21.-)
 C:Species: Homo sapiens (man)
 C:Accession: JC4795
 C:Date: 15-Oct-1995 #sequence_revision 16-Aug-1996 #text_change 19-Jul-2002
 R:Choi-Miura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Maeda, T.; Tomita, M.
 J. Biochem. 119, 1157-1165, 1996
 A:Title: Purification and characterization of a novel hyaluronan-binding protein (PHBP) activator.

A:Reference number: JC4795; MUID:96425001; PMID:8827452

A:Accession: JC4795

A:Molecule type: mRNA

A:Residues: 1-360 <CHO>

A:Cross-references: GB:S83182; NID:G1836158; PIDN:ABA46909.1; PID:G1836159

A:Experimental source: plasma

A:Note: parts of this sequence, including the amino ends of the mature chains, were detected by

C:Genetics:

A:Gene: GDB:HABP2; HARP; PHBP; HGFAL

A:Cross-references: GDB:4573962

C:Complex: a disulfide-bonded heterodimer of chains produced from the same precursor; the

C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; tryptophan

C:Keywords: chondroitin sulfate proteoglycan; glycoprotein; hyaluronic acid; hydrolase;

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-313/Product: plasma hyaluronan-binding protein, 50K chain #status predicted <50K>

F:77-108/Domain: EGF homology <EG1>

F:115-147/Domain: EGF homology <EG2>

F:154-187/Domain: EGF homology <EG3>

F:194-276/Domain: kringle homology <KRI>

F:314-550/Domain: trypsin homology <TRY>

F:314-550/Product: plasma hyaluronan-binding protein, catalytic chain #status predicted

F:54,207/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:77-88,82-97,99-108,115-125,120-136,138-147,154-165,159-176,178-187,194-276,215-257,246

F:362,405,503/Active site: His, Asp, Ser #status predicted

RESULT 18

A35005
 u-plasminogen activator (EC 3.4.21.73) precursor - chicken
 N:Alternate names: uPA
 C:Species: Gallus gallus (chicken)
 C:Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 16-Jul-1999
 C:Accession: A35005
 R:Leslie, N.D.; Kessler, C.A.; Ball, S.M.; Degen, J.L.
 J. Biol. Chem. 265, 1339-1344, 1990
 A:Title: The chicken urokinase-type plasminogen activator gene.

A:Reference number: A35005; MUID:90110185; PMID:2295632

A:Accession: A35005

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-434 <LES>

A:Cross-references: GB:J05187; NID:G212858; PIDN:AAA49131.1; PID:G212859

C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t:

C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>

F:40-71/Domain: EGF homology <EGF>

F:79-159/Domain: kringle homology <KRG>

F:173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>

F:173-416/Domain: trypsin homology <TRY>

F:162-296,202-218,210-285,310-379,342-358,369-397/Disulfide bonds: #status predicted

F:217,272,373/Active site: His, Asp, Ser #status predicted

Query Match 35.8%; Score 198.5; DB 1; Length 434;
 Best Local Similarity 54.4%; Pred. No. 1.4e-13;
 Matches 37; Conservative 7; Mismatches 19; Indels 5; Gaps 2;

QY 3 CYENGHFYRGKASTDTMGRCPLPWSNATVLTQYTHAHRSDALQLGLGKHNYCRNPDR 61

DB 79 CYSGNGEDYRGMAEDP---GCLYWDHPVIRWDYHADLKNALQLGLGKHNYCRNPDR 134

QY 62 RRPWCYVQ 69

DB 135 SRPWCYTK 142

RESULT 19

A46688
 hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
 C:Accession: A46688
 R:Wiyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
 J. Biol. Chem. 268, 10024-10028, 1993
 A:Title: Molecular cloning and sequence analysis of the cDNA for a human serine protea:

d coagulation factor XII.

A:Reference number: A46688; MUID:93252878; PMID:7683665

A:Accession: A46688

A:Molecule type: mRNA

A:Residues: 1-655 <MIY>

A:Cross-references: DDBJ:D14012; NID:G219680; PIDN:BA403113.1; PID:G219681

A:Experimental source: liver (mRNA); serum (protein)

A:Note: sequence extracted from NCBI backbone (NCBIN:131227, NCBIPI:131228)

C:Genetics:

A:Gene: GDB:HGFAC; HGFPA; HGFAP

A:Cross-references: GDB:9954514

A:Map position: 4p16-4p16

C:Function:

A:Description: activates hepatocyte growth factor by specific proteolytic cleavage

A:Pathway: tissue repair and regeneration

C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology

C:Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase

F:1-34/Domain: signal sequence #status predicted <SIG>

F:108-149/Domain: fibronectin type II repeat homology <1f2>

F:164-197/Domain: EGF homology <EG1>

F:202-237/Domain: fibronectin type I repeat homology <1f1>

F:245-278/Domain: EGF homology <EG2>

F:286-367/Domain: kringle homology <KRG>

A:Reference number: S45281; MUID:94242782; PMID:8186251

A:Accession: S45281

A:Molecule type: mRNA

A:Residues: 1-593 <SHI>

A:Cross-references: GB:S70164

A:Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue 70 as Pro, CTC for residue 203 as Phe, GTC for residue 247 as Leu, CCG for residue 286 as Cys, and ATC for residue 505 as Leu

R:Fujikawa, K.; Walsh, K.A.; Davie, E.W.

Biochemistry 16, 2270-2278, 1977

A:Title: Isolation and characterization of bovine factor XII (Hageman factor).

A:Reference number: A61329; MUID:77182112; PMID:861210

A:Accession: A61329

A:Molecule type: protein

A:Residues: 10-16, X', 18-19:525-550 <FUJ>

C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;

C:Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma;

F:37-78/Domain: fibronectin type II repeat homology <IF2>

F:88-120/Domain: EGF homology <EGF>

F:125-160/Domain: fibronectin type I repeat homology <FB1>

F:207-287/Domain: kringle homology <KRG>

F:350-587/Domain: trypsin homology <TRY>

F:541/Active site: Ser #status predicted

Query Match 28.1%; Score 155.5; DB 2; Length 593;

Best Local Similarity 38.4%; Pred. No. 7.7e-09;

Matches 33; Conservative 9; Mismatches 37; Indels 7; Gaps 3;

QY 2 TCYES--GNGHFYRGKASTDTMGSPCLPWNSATVLQQTYY-HAHRSDALQLGLGKHNYCRNP 58

DB 206 SCYDDRDRLSGVRMGAGTTLSGAPCSWAS----EATYWNVAEQVLNWLGLGHAFACRNP 261

QY 59 DNRRRPWCYVQVGLKPLVQECMVHDC 84

DB 262 DNDTRPWCPIWKGDRLSWNYCRLAPC 287

RESULT 26

A60140

plasmin (EC 3.4.21.7) precursor - chicken (fragment)

N:Alternate names: plasminogen

C:Species: Gallus gallus (chicken)

C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 16-Jul-1999

C:Accession: A60140

R:Gyenes, M.; Patthy, L.

Biochim. Biophys. Acta 832, 326-330, 1985

A:Title: The kringle 4 domain of chicken plasminogen.

A:Reference number: A60140; MUID:86077796; PMID:4074753

A:Accession: A60140

A:Molecule type: protein

A:Residues: 1-89 <GYE>

C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolo;

C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase

F:6-83/Domain: kringle homology <KRG>

F:83-27-66,55-78/Disulfide bonds: #status predicted

F:39/Binding site: carbonylate (Asn) (covalent) #status experimental

Query Match 27.6%; Score 153; DB 2; Length 89;

Best Local Similarity 38.6%; Pred. No. 2.4e-09;

Matches 34; Conservative 7; Mismatches 35; Indels 12; Gaps 4;

QY 3 CYEGNGHFYRGKASTDTMGSPCLPWNSATVLQQTYYHAHRSDALQLGLG--KHNYCRNP 60

DB 6 CYQNGSVSRGTASTFTITGKKCAWNS-----MSPRHNTKTESHFNADLRQNYCRNP 60

QY 61 RRRPWCYVQVGLKPLV--QECMVHDCAD 86

DB 61 DRSPWCYT---TDPVSRWEYCNLKRCS 85

RESULT 27

B61545

plasmin (EC 3.4.21.7) precursor - sheep (fragments)


```
Db 41 CYHNGQSYRGTSITVTGRKQOSWSMIPHRHQKTPESYPNAGLTM-----NYCENPDA 95
QY 61 RRPWCYVQVGLKPLV--QECMVHDCAD 86
Db 96 DKGPWCYT---TDPVRVWEFCNLKCKSE 120

RESULT 30
A40522
plasma (EC 3.4.21.7) precursor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 28-Feb-1992 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999
C:Accession: A40522
R:Kanalas, J.J.; Makker, S.P.
J. Biol. Chem. 266, 10825-10829, 1991
A:Title: Identification of the rat Heymann nephritis autoantigen (GP330) as a receptor s
A:Reference number: A40522; MUID:91250378; PMID:1645711
A:Accession: A40522
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-169 <KAN>
A:Cross-references: GB:M62832; NID:Q206215; PIDN:AAA41884.1; PID:Q554488
A>Note: the authors translated the codon TCT for residue 76 as Ala
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringie; serine proteinase
F:34-112/Domain: kringle homology <KRG>
F:34-112,55-95,83-107/Disulfide bonds: #status predicted

Query Match 27.1%; Score 150; DB 2; Length 169;
Best Local Similarity 34.3%; Pred.No. 9.1e-09;
Matches 37; Conservative 14; Mismatches 31; Indels 26; Gaps 7;

QY 3 CYEGNGHFYRGKASTDTMGRCPLFWNSATVLQOQYHAHRSDALQL---GLGKHNYCRNPD 59
Db 34 CYQNGKSYRGTSITVTGKQOSW-----VSMTPSHSKTTPANFPDSDL-EMNYCRNPD 87
QY 60 N-RRPWCYVQVGLKPLV--QECMVHDCAD-----GKKPSP 93
Db 88 NDQRGWCFT---TDPVRVWEFCNLKCKSETGGVAESAIVQVPSAP 132
```

Search completed: December 3, 2003, 14:44:19
Job time : 5.96552 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	554	100.0	431	1	UROK_HUMAN	P00749 homo sapien
2	519	93.7	433	1	UROK_PAPCY	P16227 papio cynoc
3	437.5	79.0	442	1	UROK_PIG	P04185 sus scrofa
4	427	77.1	432	1	UROK_RAT	P29598 rattus norv
5	432	76.2	433	1	UROK_BOVIN	Q05589 bos taurus
6	408	73.6	433	1	UROK_MOUSE	P06869 mus musculu
7	241	43.5	431	1	URTB_DESRO	P98121 desmodus ro
8	241	43.5	477	1	URTB_DESRO	P15638 desmodus ro
9	228.5	41.2	559	1	TPA_EAT	P18637 rattus norv
10	226	40.8	562	1	TPA_HUMAN	P00750 homo sapien
11	220	39.7	394	1	URTB_DESRO	P49150 desmodus ro
12	219	39.5	559	1	TPA_MOUSE	P11214 mus musculu
13	213	38.4	477	1	URTB_DESRO	P98119 desmodus ro
14	209	37.7	566	1	TPA_BOVIN	Q28198 bos taurus
15	198.5	35.8	434	1	UROK_CHICK	P15120 gallus gall
16	194.5	35.1	655	1	HGFA_HUMAN	Q04756 homo sapien
17	191.5	34.6	603	1	FA12_CAVPO	Q04962 cavia porce
18	188	33.9	653	1	HGFA_MOUSE	Q97098 mus musculu
19	170.5	30.8	615	1	FA12_HUMAN	P00748 homo sapien
20	161	29.1	1420	1	APOA_VACMU	P14417 macaca mula
21	159	28.7	473	1	KRM1_MOUSE	Q99n43 mus musculu
22	159	28.7	473	1	KRM1_RAT	Q92484 rattus norv
23	158	28.5	452	1	KRM1_YENLA	Q90y90 xenopus lae
24	157	28.3	458	1	APOA_HUMAN	P08519 homo sapien
25	156	28.2	475	1	KRM1_HUMAN	Q96mu8 homo sapien
26	155.5	28.1	593	1	FA12_BOVIN	P98140 bos taurus
27	152	27.4	812	1	PLMN_BOVIN	P06868 bos taurus
28	150	27.1	169	1	PLMN_RAT	Q01177 rattus norv
29	149	26.9	810	1	KRM2_VACMU	P12545 macaca mula
30	148.5	26.8	462	1	KRM2_HUMAN	Q8ncw0 homo sapien
31	148.5	26.8	711	1	HGFL_HUMAN	P26927 homo sapien
32	148.5	26.8	810	1	PLMN_ERIEU	Q29485 erinaceus e
33	148	26.7	333	1	PLMN_CANFA	P80009 canis famil

RA Stapleton M., Soares M.B., Ronaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Viallan D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RP [7]

RP SEQUENCE OF 66-431 FROM N.A. MEDLINE=84272706; PubMed=6589620;

RX Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.; "Identification and primary sequence of an unspliced human urokinase poly(A)+ RNA."; Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).

RL [8]

RP SEQUENCE OF 21-177. MEDLINE=83055084; PubMed=6754569;

RX Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E., Flohe L.; "The primary structure of high molecular mass urokinase from human urine. The complete amino acid sequence of the A chain."; Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).

RL [9]

RP SEQUENCE OF 156-176 AND 179-224. MEDLINE=83003608; PubMed=6749491;

RX Schaller J., Nick H., Rickli E.B., Gillesen D., Lergier W., Studer R.O.; "Human low-molecular-weight urinary urokinase. Partial characterization and preliminary sequence data of the two polypeptide chains."; Eur. J. Biochem. 125:251-257(1982).

RL [10]

RP SEQUENCE OF 158-410. MEDLINE=83055099; PubMed=6754572;

RX Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.; "The complete amino acid sequence of low molecular mass urokinase from human urine."; Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).

RL [11]

RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS). MEDLINE=96000858; PubMed=8591045;

RX Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D., Dobson C.M., Stuart D.I., Jones E.Y.; "The crystal structure of the catalytic domain of human urokinase-type plasminogen activator."; Structure 3:681-691(1995).

RL [12]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411. MEDLINE=20266327; PubMed=10805774;

RX Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G., Bode W., Magdolen V., Huber R., Moroder L.; "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly selective inhibitors of human urokinase."; Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).

RL [13]

RP STRUCTURE BY NMR. MEDLINE=89127526; PubMed=2536903;

RX Oswald R.E., Boguski M.J., Bamberger M., Smith R.A.G., Dobson C.M.; "Dynamics of the multidomain fibrinolytic protein urokinase from two-dimensional NMR."; Nature 337:579-582(1989).

RL [14]

RP STRUCTURE BY NMR OF 67-155. MEDLINE=93003110; PubMed=13271118;

RX Li X., Smith R.A.G., Dobson C.M.;

"Sequential 1H NMR assignments and secondary structure of the kringle domain from urokinase."; Biochemistry 31:9562-9571(1992).

RL [15]

RP STRUCTURE BY NMR OF 67-155. MEDLINE=94149701; PubMed=8107091;

RX Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.; "Solution structure of the kringle domain from urokinase-type plasminogen activator."; J. Mol. Biol. 235:1548-1559(1994).

RL [16]

RP VARIANT LEU-141. MEDLINE=96186279; PubMed=8652631;

RX Yoshimoto M., Ushiyama Y., Sakai M., Tanaki S., Hara H., Takahashi K., Sawasaki Y., Hanada K.; "Characterization of single chain urokinase-type plasminogen activator with a novel amino-acid substitution in the kringle structure."; Biochim. Biophys. Acta 1293:83-89(1996).

RL [17]

RP VARIANT LEU-141. MEDLINE=97218551; PubMed=9065988;

RX Conne B., Berczy M., Belin D.; "Detection of polymorphisms in the human urokinase-type plasminogen activator gene."; Thromb. Haemost. 77:434-435(1997).

RL [18]

RP ERRATUM. Conne B., Berczy M., Belin D.; Thromb. Haemost. 78:973-973(1997).

RL [19]

RP VARIANT LEU-141. MEDLINE=97337920; PubMed=9194591;

RX Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W., Creutzburg S., Graeff H., Magdolen V.; "Mutational analysis of the genes encoding urokinase-type plasminogen activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer."; Electrophoresis 18:686-689(1997).

RL CC

CC -!- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR THERAPY OF THROMBOLYTIC DISORDERS.

CC

CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.

CC

CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.

CC

CC -!- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used in Pulmonary Embolism (PE) to initiate fibrinolysis.

CC

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC

CC -!- SIMILARITY: Contains 1 kringle domain.

CC

CC -!- SIMILARITY: Contains 1 EGF-like domain.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC

DR EMBL; X02419; CAA26268.1; -

DR EMBL; M15476; AAA61253.1; -

DR EMBL; D00244; BAA00175.1; -

DR EMBL; D11143; BAA01919.1; -

DR EMBL; X02760; CAA26535.1; -

DR EMBL; AF377330; AAK53822.1; -

DR EMBL; BC013575; AAH13575.1; -

DR EMBL; K03226; AAC97138.1; -

DR EMBL; K02286; AAA61252.1; -

DR EMBL; A21571; CAA01559.1; -

DR EMBL; A18397; CAA01390.1; -

DR PIR; A00931; URGU.

DR PDB; 1KDU; 31-OCT-93.


```

Query Match      100.0%; Score 554; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 6.4e-55;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 68 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 128 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 163

RESULT 2
UROK PAPCY
ID UROK PAPCY STANDARD; PRT; 433 AA.
AC P16227;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
GN PLAU
OS Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9556;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thoracic aorta;
RX MEDLINE=90287734; PubMed=2113276;
RA A.Y.P.T., Wang T.W., Clowes A.W.;
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
RT plasminogen activator.";
RL Nucleic Acids Res. 18:3411-3411(1990).
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X51935; CAA36200.1; -
CC PIR; S14687; UKBAY.
CC HSP; P00749; 1LMW.
CC
CC MEROPS; S01.231; -
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser. protease_Try.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp spc; 1.
CC PROSITE; PS00022; EGF_1; 1.

```

```

DR PROSITE; PS01186; EGF 2; FALSE NEG.
DR PROSITE; PS00021; KRINGLE 1; 1-
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 26 62 EGF-LIKE.
FT DOMAIN 69 150 KRINGLE.
FT DOMAIN 151 177 CONNECTING PEPTIDE.
FT DOMAIN 178 433 SERINE PROTEASE.
FT DISULFID 30 38 BY SIMILARITY.
FT DISULFID 32 50 BY SIMILARITY.
FT DISULFID 52 61 BY SIMILARITY.
FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
FT DISULFID 208 224 BY SIMILARITY.
FT DISULFID 216 287 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 223 223 CHARGE RELAY SYSTEM.
FT ACT_SITE 274 274 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).
FT SEQUENCE 433 AA; 48595 MW; 816D22DFEDC8792 CRC64;

Query Match      93.7%; Score 519; DB 1; Length 433;
Best Local Similarity 94.8%; Pred. No. 5.5e-51;
Matches 91; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 67 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 126

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 127 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 162

RESULT 3
UROK PIG
ID UROK PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9923;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Kidney;
RX MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
RL Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RP REVISION TO 241.
RP Nagamine Y.;
RA Submitted (DEC-1986) to the PIR data bank.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

```

```
CC CC -!- SIMILARITY: Contains 1 kringle domain.
CC CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; X01648; CAA25806.1; -
CC CC EMBL; X02724; CAA26511.1; -
CC CC PIR; A00932; UKPG.
CC CC MEROPS; S01.231; -
CC CC InterPro; IPR001314; Chymotrypsin.
CC CC InterPro; IPR006209; EGF-like.
CC CC InterPro; IPR000001; Kringle.
CC CC InterPro; IPR001254; Ser.protease_Try.
CC CC Pfam; PF00051; kringle; 1.
CC CC Pfam; PF00089; trypsin; 1.
CC CC PRINTS; PR00722; CHYMOTRYPSIN.
CC CC PRINTS; PR00015; KRINGLE.
CC CC ProDom; PD000395; Kringle; 1.
CC CC SMART; SM00130; KR; 1.
CC CC SMART; SM00020; Tryp_SPC; 1.
CC CC PROSITE; PS00022; EGF_1; 1.
CC CC PROSITE; PS01186; EGF_2; FALSE NEG.
CC CC PROSITE; PS00021; KRINGLE_1; 1.
CC CC PROSITE; PS0070; KRINGLE_2; 1.
CC CC PROSITE; PS0240; TRYPsin_DOM; 1.
CC CC PROSITE; PS00134; TRYPsin_HIS; 1.
CC CC PROSITE; PS00135; TRYPsin_SER; 1.
CC CC KW Plasminogen activator; Hydrolase; Serine protease; Glycoprotein;
CC CC Kringle; EGF-like domain; Zymogen; Signal.
CC CC FT SIGNAL 1 20
CC CC FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
CC CC FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
CC CC FT CHAIN 190 442 CHAIN B (BY SIMILARITY).
CC CC FT DOMAIN 29 65 EGF-LIKE.
CC CC FT DOMAIN 72 153 KRINGLE.
CC CC FT DOMAIN 154 189 CONNECTING PEPTIDE.
CC CC FT DOMAIN 190 442 SERINE PROTEASE.
CC CC FT CARCCHYD 152 152 N-LINKED (GLCNAC. . .).
CC CC FT DISULFID 33 41 BY SIMILARITY.
CC CC FT DISULFID 35 53 BY SIMILARITY.
CC CC FT DISULFID 55 64 BY SIMILARITY.
CC CC FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).
CC CC FT DISULFID 220 236 BY SIMILARITY.
CC CC FT DISULFID 228 299 BY SIMILARITY.
CC CC FT DISULFID 324 393 BY SIMILARITY.
CC CC FT DISULFID 356 372 BY SIMILARITY.
CC CC FT DISULFID 383 411 BY SIMILARITY.
CC CC FT ACT_SITE 235 235 CHARGE RELAY SYSTEM.
CC CC FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
CC CC FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
CC CC FT CONFLICT 241 241 Q -> H (IN REF. 1; CAA25806).
CC CC FT CONFLICT 242 242 Q -> H (IN REF. 1; CAA26511).
CC CC FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
CC CC FT SEQUENCE 442 AA; 49116 MW; EE32FCE501321EE CRC64;
CC CC Query Match 79.0%; Score 437.5; DB 1; Length 442;
CC CC Best Local Similarity 74.3%; Pred. No. 8.4e-42;
CC CC Matches 78; Conservative 8; Mismatches 10; Indels 9; Gaps 1;
CC CC -----
CC CC 1 KTCYEGNHFYRGKASDTMGPRCLPWNASATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 60
CC CC :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC CC 70 QTCFEGNGHSGYRGKANTNTGGRPCLPWNASATVLLNTYHAHRPDALQLGLGKHNYCRNPDN 129
CC CC :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC CC 61 RRRPCYQVQGLKPLVQECMVHDC-----DGKXPSPPPE 96
CC CC :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC CC 130 QRRPCYQVQGLKPLVQECMVHDCWPCSGGSHRPAYDCKNPFSTPEK 174
CC CC :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
RESULT 4
ID UROK RAT STANDARD; PRT; 432 AA.
AC P29598;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLA0.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer 344;
RX MEDLINE=92233409; PubMed=1568219;
RA Henderson B.R.; Tansey W.P.; Phillips S.M.; Ramshaw I.A.;
RA Keiford R.F.;
RT "Transcriptional and posttranscriptional activation of urokinase
RT plasminogen activator gene expression in metastatic tumor cells.";
RL Cancer Res. 52:2489-2496(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Rabbani S.A.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH
CC CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM
CC CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE
CC LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; X63434; CAA45028.1; -
CC CC EMBL; X65651; CAA46601.1; -
CC CC PIR; S24604; S18932.
CC CC HSSP; P00749; 1KDU.
CC CC MEROPS; S01.231; -
CC CC InterPro; IPR001314; Chymotrypsin.
CC CC InterPro; IPR006209; EGF-like.
CC CC InterPro; IPR006210; EGF.
CC CC InterPro; IPR000001; Kringle.
CC CC InterPro; IPR001254; Ser.protease_Try.
CC CC Pfam; PF00051; kringle; 1.
CC CC Pfam; PF00089; trypsin; 1.
CC CC PRINTS; PR00722; CHYMOTRYPSIN.
CC CC PRINTS; PR00015; KRINGLE.
CC CC ProDom; PD000395; Kringle; 1.
CC CC SMART; SM00181; EGF; 1.
CC CC SMART; SM00130; KR; 1.
CC CC SMART; SM00020; Tryp_SPC; 1.
CC CC PROSITE; PS00022; EGF_1; 1.
CC CC PROSITE; PS01186; EGF_2; FALSE NEG.
CC CC PROSITE; PS00021; KRINGLE_1; 1.
CC CC PROSITE; PS00020; KRINGLE_2; 1.
CC CC PROSITE; PS0240; TRYPsin_DOM; 1.
CC CC PROSITE; PS00134; TRYPsin_HIS; FALSE NEG.
CC CC PROSITE; PS00135; TRYPsin_SER; 1.
```

```

KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 20 177 CHAIN A (BY SIMILARITY).
FT CHAIN 156 177 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 179 432 CHAIN B (BY SIMILARITY).
FT DOMAIN 27 63 EGF-LIKE.
FT DOMAIN 70 151 KRINGLE.
FT DOMAIN 152 178 CONNECTING PEPTIDE.
FT DOMAIN 179 432 SERINE PROTEASE.
FT DISULFID 31 39 BY SIMILARITY.
FT DISULFID 33 51 BY SIMILARITY.
FT DISULFID 53 62 BY SIMILARITY.
FT DISULFID 168 300 INTERCHAIN (BY SIMILARITY).
FT DISULFID 210 226 BY SIMILARITY.
FT DISULFID 218 269 BY SIMILARITY.
FT DISULFID 314 383 BY SIMILARITY.
FT DISULFID 346 362 BY SIMILARITY.
FT DISULFID 373 401 BY SIMILARITY.
FT ACT_SITE 225 225 CHARGE RELAY SYSTEM.
FT ACT_SITE 276 276 CHARGE RELAY SYSTEM.
FT ACT_SITE 377 377 CHARGE RELAY SYSTEM.
FT CONFLICT 16 16 N -> H (IN REF. 2).
FT CONFLICT 24 24 E -> G (IN REF. 2).
FT CONFLICT 332 332 D -> N (IN REF. 2).
SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;

Query Match 77.1%; Score 427; DB 1; Length 432;
Best Local Similarity 77.1%; Pred. No. 1.3e-40;
Matches 74; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPDN 60
DB 68 KTCYHGNGSGYRGKANTDTKGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPDN 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKXPSSPPEE 96
DB 128 QRAPWCYVQVGLKPLVQECMVHDCADGKXPSSPPEE 163

RESULT 5
ID UROK BOVIN STANDARD; PRT; 433 AA.
AC Q05589; Q28209.
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Bos taurus (Bovine).
OC Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aortic endothelium;
RX MEDLINE=93216119; PubMed=8385052;
RA Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,
RA Schleuning W.-D.;
RT "Bovine urokinase-type plasminogen activator and its receptor:
RT cloning and induction by retinoic acid.";
RL Gene 125:177-183(1993).
RN [2]
RP SEQUENCE OF 12-433 FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Bejrgund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RT and tPA.";
RL Int. Dairy J. 5:605-617(1995).
DB -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in

```

```

CC plasminogen to form plasmin.
CC -1- INDUCTION: By retinoic acid.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L03546; AA051419.1; -.
DR EMBL; X85801; CA059796.1; -.
DR PIR; JN0560; JN0560.
DR HSP; P00749; ILMW.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 180 CONNECTING PEPTIDE.
FT DOMAIN 181 433 SERINE PROTEASE.
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 189 189 A -> T (IN REF. 2).
SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;

Query Match 76.2%; Score 422; DB 1; Length 433;
Best Local Similarity 75.0%; Pred. No. 4.6e-40;
Matches 72; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPDN 60
DB 70 KTCYHGNGSGYRGKANDLGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPDN 129
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKXPSSPPEE 96
DB 130 QRAPWCYVQVGLKPLVQECMVHDCADGKXPSSPPEE 165

```

```
RESULT 6
UROK_MOUSE
ID UROK_MOUSE STANDARD; PRT; 433 AA.
AC P06859;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. PubMed=2985383;
RX MEDLINE=85179474;
RA Belin D., Vassalli J.-D., Combepine C., Godeau F., Nagamine Y.,
RA Reich E., Kocher H.P., Duvoisin R.M.;
RT "Cloning, nucleotide sequencing and expression of cDNAs encoding
RT mouse urokinase-type plasminogen activator.";
RL Eur. J. Biochem. 148:225-232(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88163489; PubMed=2831940;
RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;
RT "The murine urokinase-type plasminogen activator gene.";
RL Biochemistry 26:8270-8279(1987).
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X02389; CAA26231.1; -
CC EMBL; M17922; AAA40539.1; -
CC PIR; A29420; UKMS.
CC HSP; P00749; 1KDU.
CC MEROPS; S01.231; -.
CC MGD; MGI.97611; Plau.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser. protease_Try.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC ProDom; SM00181; EGF; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp. SEC; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; FALSE_NEG.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PSS0070; KRINGLE_2; 1.
CC PROSITE; PSS0240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
CC
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 178 CHAIN A (BY SIMILARITY).
FT CHAIN 157 178 SHORT A CHAIN (A1).
FT CHAIN 180 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 28 64 EGF-LIKE.
FT DOMAIN 71 152 KRINGLE.
FT DOMAIN 153 179 CONNECTING PEPTIDE.
FT DOMAIN 180 433 SERINE PROTEASE.
FT DISULFID 32 40 BY SIMILARITY.
FT DISULFID 34 52 BY SIMILARITY.
FT DISULFID 54 63 BY SIMILARITY.
FT DISULFID 169 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM.
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
SQ SEQUENCE 433 AA; 48268 MW; A99C35F6250443F9 CRC64;
Query Match 73.6%; Score 408; DB 1; Length 433;
Best Local Similarity 70.8%; Pred No. 1,7e-38;
Matches 68; Conservative 12; Mismatches 16; Indels 0; Gaps 0;
QY 1 KTCYEGNHFVGRGKASTDTMGRPCLPWNSATVLTQTYHAHRSALQLGLGKHNYCRNPDN 60
DB 69 KTCYHGNGSDYRGKANTDTTKGRPCLANAPALQKPNYNAHRPDALSLGLGKHNYCRNPDN 128
QY 61 RRRPCWYQVGLKPLVQECWYHDCADGKKPSPEE 96
DB 129 QRRPCWYQVGLRQFVQECWYHDCSLSKKPSVVQ 164
RESULT 7
URTB DESRO STANDARD; PRT; 431 AA.
ID URTB DESRO
AC P98121;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA
DE beta).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
```



```
DR EMBL: J05082; AAA31596.1; -.
DR PIR: J05098; J05098.
DR HSP: P98119; IAS1.
DR MEROPS: S01.232; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR006209; EGF-like.
DR InterPro: IPR000083; Fibronectin.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00039; fnl; 1.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; kringle; 1.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00038; fnl; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS01253; FIBRONECTIN_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS02040; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR KEGG: Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 477
FT DOMAIN 40 82
FT DOMAIN 83 121
FT DOMAIN 128 209
FT DOMAIN 225 477
FT ACT_SITE 272 272
FT ACT_SITE 321 321
FT ACT_SITE 428 428
FT DISULFID 42 72
FT DISULFID 70 79
FT DISULFID 87 98
FT DISULFID 92 109
FT DISULFID 111 120
FT DISULFID 128 209
FT DISULFID 149 191
FT DISULFID 180 204
FT DISULFID 214 345
FT DISULFID 257 273
FT DISULFID 265 334
FT DISULFID 359 434
FT DISULFID 391 407
FT DISULFID 424 452
FT CARBOHYD 185 185
FT CARBOHYD 398 398
FT CONFLICT 403 403
FT CONFLICT 417 417
FT CONFLICT 435 435
FT SEQUENCE 477 AA; 53719 MW; 17486555C085077C CRC64;
Query Match 43.5%; Score 241; DB 1; Length 477;
Best Local Similarity 50.0%; Pred. No. 1.2e-19;
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;
QY 2 TCYEGNGHYGRKASTDTMGRCLPWNSTVLQTYHARSALQGLGKHNYCRPNDR 61
Db 127 TCYKQGVYRGVTSWTSSESQAQCNWNSLLTRTYNGRRSDAITLGLGKHNYCRPNDR 186
QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
Db 187 SKPCYVIRASKILFCSVPVCS 210
```

RESULT 9

TPA RAT

ID TPA RAT STANDARD; PRT; 559 AA.

AC P19637;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-NOV-1991 (Rel. 20, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)

DE (t-PA) (t-plasminogen activator).

GN PLAT.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

SEQUENCE FROM N.A.

RP MEDLINE=89170114; PubMed=3148445;

RA Ny T.; Leonardsson G.; Hsueh A.J.W.;

RT "Cloning and characterization of a cDNA for rat tissue-type

RT plasminogen activator.";

RL DNA 7:671-677(1988).

RN [2]

SEQUENCE FROM N.A.

RP MEDLINE=90130448; PubMed=2105315;

RA Feng P.; Ohlsson M.; Ny T.;

RT "The structure of the TATA-less rat tissue-type plasminogen activator

RT gene. Species-specific sequence divergences in the promoter predict

RT differences in regulation of gene expression.";

RL J. Biol. Chem. 265:2022-2027(1990).

CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN

CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY

CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT

CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND

CC MANY OTHER PHYSIOLOGICAL EVENTS.

CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in

CC plasminogen to form plasmin.

CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE

CC BOND.

CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.

CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME CAN BE FURTHER

CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER

CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.

CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A

CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -!- SIMILARITY: Contains 1 EGF-like domain.

CC -!- SIMILARITY: Contains 1 fibronectin type I domain.

CC -!- SIMILARITY: Contains 2 kringle domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL: M23697; AAA41812.1; -

DR EMBL: M31197; AAA42261.1; -

DR EMBL: M31185; AAA42261.1; JOINED.

DR EMBL: M31186; AAA42261.1; JOINED.

DR EMBL: M31187; AAA42261.1; JOINED.

DR EMBL: M31188; AAA42261.1; JOINED.

DR EMBL: M31189; AAA42261.1; JOINED.

DR EMBL: M31190; AAA42261.1; JOINED.

DR EMBL: M31191; AAA42261.1; JOINED.

DR EMBL: M31192; AAA42261.1; JOINED.

DR EMBL: M31193; AAA42261.1; JOINED.

DR EMBL: M31194; AAA42261.1; JOINED.

DR EMBL: M31195; AAA42261.1; JOINED.

DR EMBL: M31196; AAA42261.1; JOINED.

DR EMBL: A19618; CAA01482.1; -

```
DR PIR: A35029; A35029.
DR HSP: P00750; 1RTE.
DR MEROPS: S01.232; .
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR006209; EGF-like.
DR InterPro: IPR000083; Fibrinectin.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00039; fn1; 1.
DR Pfam: PF00051; kringle; 2.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PRO0722; CHYMOTRYPSIN.
DR PRINTS: PRO0018; KRINGLE.
DR ProDom: PD000395; Kringle; 2.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00058; fn1; 1.
DR SMART: SM00130; KR; 2.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS01253; FIBRONECTIN_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 2.
DR PROSITE: PS00070; KRINGLE_2; 2.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR Plasma; Kringle; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17 PROBABLE.
FT PROPEP 18 29
FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT CHAIN 36 78 FIBRONECTIN TYPE-1.
FT DOMAIN 79 117 EGF-LIKE.
FT DOMAIN 124 205 KRINGLE 1.
FT DOMAIN 213 294 KRINGLE 2.
FT DOMAIN 309 559 SERINE PROTEASE.
FT ACT_SITE 355 355 CHARGE RELAY SYSTEM.
FT ACT_SITE 404 404 CHARGE RELAY SYSTEM.
FT ACT_SITE 510 510 CHARGE RELAY SYSTEM.
FT DISULFID 38 68 BY SIMILARITY.
FT DISULFID 66 75 BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 294 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 348 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 380 380 E -> K (IN REF. 1).
SQ SEQUENCE 559 AA; 62903 MW; 7DBD3809C1D1C921 CRC64;
Query Match 41.2%; Score 228.5; DB 1; Length 559;
Best Local Similarity 45.3%; Pred. No. 3.6e-18;
Matches 43; Conservative 11; Mismatches 36; Indels 5; Gaps 1;
DR 2 TCVEGNHGYRGKASPTDWTGRPCLPNWSATVLQOTVHAHRSALQLGLGKHNCRPNDR 61
123 TCPEGQGITVGTWSTAENGAECINMSSALSQKPSARRPNAIKLGLGNHNYCRNPRD 182
62 RRPWCYVQGLKPLVQECMVHDCADGKKPSSPPEE 96
183 VKPWCYVFKAGYTTTEFCSTPAC-----PKGPTED 212
RESULT 10
TPA_HUMAN STANDARD; PRT; 562 AA.
AC P00750; Q15103; .
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tissue-type plasminogen activator precursor (BC 3.4.21.68) (tpa)
DE (t-PA) (t-plasminogen activator) (Alteplase) (Retepase).
GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE-Melanoma;
RX MEDLINE=83115262; PubMed=6337343;
RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA Goeddel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RT cDNA in E. coli.";
RL Nature 301:214-221(1983).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=88262579; PubMed=3133640;
RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT from human fetal lung cells.";
RL Nucleic Acids Res. 16:5695-5695(1988).
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=88054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,
RA Hsiung N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RT mouse cells using BPV vectors.";
RL DNA 6:461-472(1987).
[4]
RN SEQUENCE FROM N.A.
RX MEDLINE=86196143; PubMed=3009482;
RA Frieze Degen S.J., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene.";
J. Biol. Chem. 261:6972-6985(1986).
[5]
RN SEQUENCE FROM N.A.
RX MEDLINE=84298137; PubMed=6089198;
RA Ny T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RT correlation of intron and exon structures to functional and
RT structural domains.";
Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
[6]
RN SEQUENCE FROM N.A.
RX MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Entage J.S.,
RA Opdenakker G., Volckaert G., Rombauts W., Billiau A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RT and its expression in Escherichia coli.";
Mol. Biol. Med. 3:279-292(1986).
[7]
RN SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Unbilical vein;
RX MEDLINE=90192129; PubMed=2107528;
```

RA Siebert P.D., Fong K.;
 RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
 RT human endothelial cells";
 RL Nucleic Acids Res. 18:1086-1086 (1990).
 RN (8)
 RN SEQUENCE FROM N.A.
 RP TISSUE=Brain;
 RC MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M.J., Udell T.B., Toshikiyuki S., Carninci P., Frange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smalusz D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Maria M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN (9)
 RN SEQUENCE OF 212-361 FROM N.A.
 RP MEDLINE=8316656; PubMed=6572897;
 RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,
 RA Josephson S.;
 RT "Isolation of cDNA sequences coding for a part of human tissue
 RT plasminogen activator";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352 (1983).
 RN (10)
 RN SEQUENCE OF 1-36 FROM N.A.
 RP MEDLINE=85289338; PubMed=1611993;
 RA Fisher R., Waller E.K., Grossi R., Thompson D., Tizard R.,
 RA Schleuning W.-D.;
 RT "Isolation and characterization of the human tissue-type plasminogen
 RT activator structural gene including its 5' flanking region";
 RL J. Biol. Chem. 260:11223-11230 (1985).
 RN (11)
 RN SEQUENCE OF 31-562 FROM N.A.
 RP MEDLINE=91293340; PubMed=1368681;
 RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
 RT "Purification and characterization of tissue plasminogen activator
 RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells";
 RL Agric. Biol. Chem. 55:1225-1232 (1991).
 RN (12)
 RN SEQUENCE OF 36-562.
 RC TISSUE=Melanoma;
 RX MEDLINE=85000468; PubMed=6433976;
 RA Pohl G., Kallstroem M., Bergsdorf N., Wallen P., Joernvall H.;
 RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
 RT derived amino acid sequence, identify the active site serine residue,
 RT establish glycosylation sites, and localize variant differences";
 RL Biochemistry 23:3701-3707 (1984).
 RN (13)
 RN SEQUENCE OF 33-52 AND 311-330.
 RC TISSUE=Melanoma;
 RX MEDLINE=83209620; PubMed=6682760;
 RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;
 RT "Purification and characterization of a melanoma cell plasminogen
 RT activator";
 RL Eur. J. Biochem. 132:681-686 (1983).
 RN (14)
 RN STRUCTURE OF CARBOHYDRATES.
 RP MEDLINE=90092112; PubMed=2513186;
 RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
 RT "Carbohydrate structure of recombinant human uterine tissue

RT plasminogen activator expressed in mouse epithelial cells";
 RL Eur. J. Biochem. 186:273-286 (1989).
 RN (15)
 RP CARBOHYDRATE-LINKAGE SITE THR-96.
 RX MEDLINE=91159408; PubMed=1900431;
 RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
 RT "Tissue plasminogen activator has an O-linked fucose attached to
 RT threonine-61 in the epidermal growth factor domain";
 RL Biochemistry 30:2311-2314 (1991).
 RN (16)
 RN DISULFIDE BONDS IN KRINGLE 2.
 RP MEDLINE=91244765; PubMed=1645336;
 RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;
 RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
 RT plasminogen activator produced in *Escherichia coli*";
 RL J. Biol. Chem. 266:10070-10072 (1991).
 RN (17)
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=96200985; PubMed=8613982;
 RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
 RA Bode W.;
 RT "The 2.3 Å crystal structure of the catalytic domain of recombinant
 RT two-chain human tissue-type plasminogen activator";
 RL J. Mol. Biol. 258:117-135 (1996).
 RN (18)
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
 RX MEDLINE=97449126; PubMed=9305622;
 RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
 RA Bode W.;
 RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
 RT crystal structure of single-chain human tPA";
 RL EMBO J. 16:4797-4805 (1997).
 RN (19)
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
 RX MEDLINE=92118803; PubMed=1310033;
 RA de Vos A., Uitsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,
 RA Westbrook M.L., Kosiakof A.A.;
 RT "Crystal structure of the kringle 2 domain of tissue plasminogen
 RT activator at 2.4 Å resolution";
 RL Biochemistry 31:270-279 (1992).
 RN (20)
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=90122799; PubMed=2558718;
 RA Byeon I.-J.L., Kelley R.F., Llinas M.;
 RT "1H NMR structural characterization of a recombinant kringle 2 domain
 RT from human tissue-type plasminogen activator";
 RL Biochemistry 28:9350-9360 (1989).
 RN (21)
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=91200042; PubMed=1901789;
 RA Byeon I.-J.L., Kelley R.F., Llinas M.;
 RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR
 RT assignments and secondary structure";
 RL Eur. J. Biochem. 197:155-165 (1991).
 RN (22)
 RP STRUCTURE BY NMR OF KRINGLE 2.
 RX MEDLINE=92106329; PubMed=1762144;
 RA Byeon I.-J.L., Llinas M.;
 RT "Solution structure of the tissue-type plasminogen activator kringle
 RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
 RT drug";
 RL J. Mol. Biol. 222:1035-1051 (1991).
 RN (23)

Query Match 40.8%; Score 226; DB 1; Length 562;
 Best Local Similarity 47.7%; Pred. No. 6.9e-18;
 Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSTVLOQTYHAHRS DALQLGLGHNYCRPNDR 61
 Db 126 TCYEDQGISYAGTWSTAESAECTNWS SALLAKPYSGRPDAIRLGLGHNYCRPNDR 185
 QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87

KW	Kringle; Signal; Multigene family.
FT	SIGNAL 1 36 POTENTIAL.
FT	CHAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
FT	DOMAIN 45 126 KRINGLE.
FT	DOMAIN 142 394 SERINE PROTEASE.
FT	ACT_SITE 189 189 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 345 345 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID 45 126 BY SIMILARITY.
FT	DISULFID 66 108 BY SIMILARITY.
FT	DISULFID 97 121 BY SIMILARITY.
FT	DISULFID 131 262 BY SIMILARITY.
FT	DISULFID 174 190 BY SIMILARITY.
FT	DISULFID 182 251 BY SIMILARITY.
FT	DISULFID 276 351 BY SIMILARITY.
FT	DISULFID 308 324 BY SIMILARITY.
FT	DISULFID 341 369 BY SIMILARITY.
FT	CARBOHYD 315 315 N-LINKED (GLCNAC..) (POTENTIAL).
QY	SEQUENCE 394 AA; 44105 MW; 9CDD6F52F3D81 FCD CRC64;
QY	Query Match 39.78; Score 220; DB 1; Length 394;
QY	Best Local Similarity 45.2; Pred.No.2.3e-17;
QY	Matches 38; Conservative 13; Mismatches 33; Indels 0; Gaps 0
Db	2 TCYEGNGHGFYRGKASTDMGRPCLPWRKSATVLOOTYYHAHRSDALQLGGLGHKNYCRNPDR 61 : : : : : : : : : : :
Db	44 TCYKDQGVTYRGTWSTSEGAQCINMNSLLIRTYNGRMPEAVKGLGNHNHCNPDGA 103 : : : : : : : : : : :
Qy	62 RRPWCYVOVLKPLVQECMVHDCA 85 : : : :
Db	104 SKPCYVIKARKFTSCECVFVC 127 : : : :
RESULT 12	
TPA_MOUSE	
ID	TPA_MOUSE STANDARD; PRT; 559 AA.
AC	P1214; Q91VP2;
DT	01-JUN-1989 (Rel. 11, Created)
DT	15-SEP-2003 (Rel. 42, Last sequence update)
DT	15-SEP-2003 (Rel. 42, Last annotation update)
DE	Tissue-type plasminogen activator precursor (EC 3.4.21.66) (tPA)
DE	(t-PA) (t-plasminogen activator).
GN	PLAT.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NCBI_TaxID	10090;
RN	[1]
RN	SEQUENCE FROM N.A.
RP	MEDLINE=8608703; PubMed=2826484;
RX	Medline=22388257; Pubmed=12477932;
RA	Rickles R.J., Darrow A.L., Strickland S.;
RA	"Molecular cloning of complementary DNA to mouse tissue plasminogen
RA	activator mRNA and its expression during F9 teratocarcinoma cell
RA	differentiation";
RL	J Biol. Chem. 263:1563-1569(1988).
RN	[2]
RN	SEQUENCE FROM N.A.
RP	TSSU=Mammary gland;
RX	MEDLINE=22388257; PubMed=12477932;
RA	Strauszberg R.L., Feingold E.A., Schuler G.D.,
RA	Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,
RA	Alschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marasina K., Farmer A.A., Rubin G.W., Hong L.,
RA	Scapellato M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA	Bosa S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villation D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Faney J., Helton E., Kettenen M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.J., Skalska U., Smallos D.E.,
RA Schenrich A., Schein J.E., Jones S.J.M., Marra M.A.:
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -1- PM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-108 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 fibronectin type 1 domain.
CC -1- SIMILARITY: Contains 2 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J03520; AAA0470.1; --
CC EMBL: BC011256; AAH1256.1; --
CC PIR: A29941; A29941.
CC HSP: P00750; 1A5H.
CC MEROPS: S01.232; --
CC MGD: MGI:97610; Plat.
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR006209; EGF-like.
CC InterPro: IPR000083; Fibritncl.
CC InterPro: IPR006210; IEGF.
CC InterPro: IPR000001; Kringle.
CC InterPro: IPR001254; Ser_protease_try.
CC Pfam: PF00008; EGF_1.
CC Pfam: PF00039; fn1_1.
CC Pfam: PF00051; kringle_2.
CC Pfam: PF00089; trypsin_1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC PRINTS: PR00018; KRINGLE.
CC ProDom: PD000395; Kringle_2.
CC SMART: SM00181; EGF_1.
CC SMART: SM00058; FN1_1.
CC SMART: SM00130; KR_2.
CC SMART: SM00020; Tryp_Spc_1.
CC PROSITE: PS00022; EGF_1; 1.
CC PROSITE: PS01186; EGF_2; 1.
CC PROSITE: PS01253; FIBRONECTIN_1; 1.
CC PROSITE: PS00021; KRINGLE_1; 2.
CC PROSITE: PS00070; KRINGLE_2; 2.
CC PROSITE: PS02040; TRYPSIN_DOM; 1.
CC PROSITE: PS00134; TRYPSIN_HIS; 1.
CC PROSITE: PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17 PROBABLE.
FT PROPEP 18 29
FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B

FT DOMAIN 36 78
FT FIBRONECTIN TYPE-1.
FT EGF-LIKE
FT KRINGLE 1.
FT KRINGLE 2.
FT SERINE PROTEASE.
FT ACT SITE 355 355
FT CHARGE RELAY SYSTEM.
FT ACT SITE 404 404
FT CHARGE RELAY SYSTEM.
FT DISULFID 38 68
FT BY SIMILARITY.
FT DISULFID 66 75
FT BY SIMILARITY.
FT DISULFID 83 94
FT BY SIMILARITY.
FT DISULFID 88 105
FT BY SIMILARITY.
FT DISULFID 107 116
FT BY SIMILARITY.
FT DISULFID 124 205
FT BY SIMILARITY.
FT DISULFID 145 187
FT BY SIMILARITY.
FT DISULFID 176 200
FT BY SIMILARITY.
FT DISULFID 213 294
FT BY SIMILARITY.
FT DISULFID 234 276
FT BY SIMILARITY.
FT DISULFID 265 289
FT BY SIMILARITY.
FT DISULFID 297 428
FT BY SIMILARITY.
FT DISULFID 340 356
FT BY SIMILARITY.
FT DISULFID 348 417
FT BY SIMILARITY.
FT DISULFID 442 516
FT BY SIMILARITY.
FT DISULFID 474 490
FT BY SIMILARITY.
FT DISULFID 506 534
FT BY SIMILARITY.
FT CARBOHYD 149 149
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 260 260
FT G -> A (IN REF. 1).
FT CONFLICT 325 325
FT P -> A (IN REF. 1).
SQ SEQUENCE 559 AA; 63122 MW; 8CCE52BDB94514D9 CRC64;

Query Match 39.5%; Score 219; DB 1; Length 559;
Best Local Similarity 46.0%; Pred. No. 4.2e-17;
Matches 40; Conservative 11; Mismatches 36; Idels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRLPWSATVLOQTYHAHRSALQLGLGHNYCRNPDNR 61
DB 123 TCFEQGITRYGTWSTAESGAECINWSSVLSPYNARRPNAIKLGLGHNYCRNPDNR 182
QY 62 RRPWCYVQVGLKPLVQECWHDGADGK 88
DB 183 LRPCWYVFKAGRYTTEFCSTACPKGK 209

RESULT 13
ID _UTL1_DESRO STANDARD; PRT; 477 AA.
AC P98119; 1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSPA
DE alpha-1).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RA "The plasminogen activator family from the salivary gland of the
RA vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,

DB 127 TCYEGQGVYRTGTWSTABSRVEGINWSSLLTRTYNGRMDFAFNLGLGNHNYCRNPNGA 186
QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
DB 187 KPPWCYVYKAGKFTSSECSVPVCS 210

RESULT 14
TPA_BOVIN STANDARD; PRT; 566 AA.
ID TPA_BOVIN
AC Q26198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
and tPA.";
RL Int. Dairy J. 5:605-617(1995).
CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
MANY OTHER PHYSIOLOGICAL EVENTS.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
BOND.
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 2 kringle domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch.
CC
CC -----
CC EMBL; X85800; CAA59795.1; -;
CC HSP; P00750; IRTF.
CC MEROPS; S01.232; -;
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006203; EGF-like.
CC InterPro; IPR000083; Fibnctnrl.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00039; fnl; 1.
CC Pfam; PF00051; kringle; 2.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 2.

DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FNI; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS0070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRIPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Plasminogen activator; Hydrolase; Serine protease; Glycoprotein;
KW plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 21
FT PROPEP 22 33 BY SIMILARITY.
FT CHAIN 34 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 34 314 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
CHAIN.
FT CHAIN 315 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
CHAIN.
FT DOMAIN 40 82 FIBRONECTIN TYPE-1.
FT DOMAIN 83 121 EGF-LIKE.
FT DOMAIN 128 209 KRINGLE 1.
FT DOMAIN 219 300 KRINGLE 2.
FT DOMAIN 315 566 SERINE PROTEASE.
FT ACT_SITE 361 361 CHARGE RELAY SYSTEM.
FT ACT_SITE 410 410 CHARGE RELAY SYSTEM.
FT ACT_SITE 517 517 CHARGE RELAY SYSTEM.
FT DISULFID 72 72 BY SIMILARITY.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 87 98 BY SIMILARITY.
FT DISULFID 92 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 128 209 BY SIMILARITY.
FT DISULFID 149 191 BY SIMILARITY.
FT DISULFID 180 204 BY SIMILARITY.
FT DISULFID 219 300 BY SIMILARITY.
FT DISULFID 240 282 BY SIMILARITY.
FT DISULFID 271 295 BY SIMILARITY.
FT DISULFID 303 434 INTERCHAIN (BY SIMILARITY).
FT DISULFID 346 362 BY SIMILARITY.
FT DISULFID 354 423 BY SIMILARITY.
FT DISULFID 448 523 BY SIMILARITY.
FT DISULFID 480 496 BY SIMILARITY.
FT DISULFID 513 541 BY SIMILARITY.
FT CARBOHYD 153 153 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 487 487 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 566 AA; 63701 MW; 2EB6BEB4E32276C3 CRC64;
Query Match 37.7%; Score 209; DB 1; Length 566;
Best Local Similarity 47.6%; Pred. No. 5.7e-16;
Matches 39; Conservative 9; Mismatches 34; Indels 0; Gaps 0;
QY 3 CYEGNGHYRFGKASDTMTGRPCLPWNSATVLQQTVAHRSDALQLGLGKHNYCRNPNNR 62
DB 219 CYTGNGLAYRGRSHRTKSGASCLPWNVSFLTISKIYTAKWSNAPALGLGKHNCNPDDGA 278
QY 63 RPPWCYVQVGLKPLVQECMVHDC 84
DB 279 QPWCYVQVGLKPLVQECMVHDC 300
PRT; 434 AA.
RESULT 15
UROK_CHICK
ID UROK_CHICK STANDARD; PRT; 434 AA.
AC P15120;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RX MEDLINE=90110185; PubMed=2295632;
RA Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;
RT "The chicken urokinase-type plasminogen activator gene";
RL J. Biol. Chem. 265:1339-1344 (1990).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J05187; AAA49131.1; -
DR EMBL; J05188; AAA49130.1; -
DR PIR; A35005; A35005.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.231; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00049; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 434
FT CHAIN 21 171
FT CHAIN 173 434
FT DOMAIN 36 72
FT DOMAIN 79 158
FT DOMAIN 159 172
FT DOMAIN 173 434
FT DISULFID 40 48
FT DISULFID 42 60
FT DISULFID 62 71
FT DISULFID 162 296
FT DISULFID 202 218
FT DISULFID 210 285
FT DISULFID 310 379
FT DISULFID 342 358
FT DISULFID 369 397
FT ACT_SITE 217 272
FT ACT_SITE 272 272
FT ACT_SITE 373 373
FT CARBOHYD 228 228
SQ SEQUENCE 434 AA; 49400 MW; BD881048DD866A55 CRC64;

Query Match 35.8%; Score 198.5; DB 1; Length 434;
Best Local Similarity 54.4%; Pred. No. 6.6e-15; Indels 5; Gaps 2;
Matches 37; Conservative 7; Mismatches 19;
QY 3 CYEGNGHYRGKASTDTMGRCPLPWSATVLQ-QTYHAHRSDALQLGLGKHYCRNPGR 61
DB 79 CYSNGEDYRGMAEDP----GCLYWDHPSVIRMGDYHADLKNALQLGLGKHYCRNPGR 134
QY 62 RRPWCYVQ 69
DB 135 SRPWCYTK 142
RESULT 16
HGFA_HUMAN STANDARD; PRT; 655 AA.
AC Q04756; Q14726;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFAc.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver, and Serum;
RX MEDLINE=91252878; PubMed=7683665;
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
RA Kitamura N.;
RT "Molecular cloning and sequence analysis of the cDNA for a human
RT serine protease responsible for activation of hepatocyte growth
RT factor. Structural similarity of the protease precursor to blood
RT coagulation factor XII.";
RL J. Biol. Chem. 268:10024-10028 (1993).
RN [2]
RP SEQUENCE OF 40-655 FROM N.A.
RA Zhao S., Odell C.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY
CC CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
CC -1- TISSUE SPECIFICITY: LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D14012; BAA03113.1; -
DR EMBL; Z69923; CAA93803.1; -
DR PIR; A46688; A46688.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.228; -
DR Genew; HGNC:4894; HGFAc.
DR MIM; 604552; -
DR GO; GO:0005576; C:extracellular; TAS.


```

RESULT 20
APOA MACMU
ID APOA MACMU STANDARD; PRT; 1420 AA.
AC P14417.
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Apolipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Lp(a)) (Fragment).
GN LPA.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89174660; PubMed=2925643;
RA Tomlinson J.E., McLean J.W., Lawn R.M.;
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
synthesis.";
RL J. Biol. Chem. 264:5957-5965(1989).
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
(Lp(a)). It has serine proteinase activity and is able of
autolysis. Inhibits tissue-type plasminogen activator 1.
Lp(a) may be a ligand for megalin/Gp 330.
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
decorin (by similarity).
CC -!- PTM: N- and O-glycosylated (by similarity).
CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its
naturally occurring proteolytic fragments are correlated with
atherosclerosis. Homology with plasminogen kringle IV and V is
thought to underlie the atherogenicity of the protein, because the
fragments are competing with plasminogen for fibrinogen binding.
CC -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
leading to the formation of the so called mini-Lp(a). Apo(a)
fragments accumulate in atherosclerotic lesions, where they may
promote thrombogenesis. O-glycosylation may limit the extent of
proteolytic fragmentation (By similarity).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -!- SIMILARITY: Contains at least 10 kringle domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; J04635; AAA36833.1; -.
CC PIR; A32869; A32869.
CC HSP; P00747; 2PK4.
CC MEROPS; S01-226; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Serine protease_Try.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 10.
CC SMART; SM00330; KR; 10.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00021; KRINGLE 1; 10.
CC PROSITE; PS00070; KRINGLE 2; 10.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
CC PROSITE; PS00135; TRYPSIN_SER; FALSE_NEG.
CC Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
KW Kringle; Repea; 1
FT NON_TER 1
FT DOMAIN 49 127 KRINGLE 1.
FT DOMAIN 163 241 KRINGLE 2.

```

```

FT DOMAIN 277 355 KRINGLE 3.
FT DOMAIN 391 469 KRINGLE 4.
FT DOMAIN 505 583 KRINGLE 5.
FT DOMAIN 619 697 KRINGLE 6.
FT DOMAIN 725 803 KRINGLE 7.
FT DOMAIN 839 917 KRINGLE 8.
FT DOMAIN 953 1031 KRINGLE 9.
FT DOMAIN 1067 1145 KRINGLE 10.
FT DOMAIN 1191 1420 SERINE PROTEASE.
SQ SEQUENCE 1420 AA; 158367 MW; B8102949E03C5B0E CRC64;

Query Match 29.1%; Score 161; DB 1; Length 1420;
Best Local Similarity 40.9%; Pred No. 3.8e-10;
Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;

QY 3 CYEGNGHFYRGKASTDMGRPCLPWN SATVLQ--QTYHAHRSDALQLGLGKHYCRNPN 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1068 CYHNGGYSQYGFSTVTGRTCSWSSMTPHQKRTPENHPNDLTM-----NYCRNPDA 1122

QY 61 RRRPCYVQVGLKPLVQE--CMVHDCAD 86
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1123 DTGFWCFT---MDPSVRREYCNLTGCS 1147

RESULT 21
KRM1 MOUSE
ID KRM1 MOUSE STANDARD; PRT; 473 AA.
AC Q99M43.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringle-containing protein marking the eye
and the nose) (Dickkopf receptor).
GN KREMEN1 OR KREMEN
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., DEVELOPMENTAL STAGE, AND TISSUE SPECIFICITY.
RX TISSUE=Brain, and Kidney; PubMed=11267660;
RX MEDLINE=21167372; PubMed=11267660;
RA Nakamura T., Aoki S., Kitajima K., Takahashi T., Matsumoto K.,
RA Nakamura T.;
RT "Molecular cloning and characterization of Kremen, a novel
Kringle-containing transmembrane protein.";
RL Biochim. Biophys. Acta 1518:63-72(2001).
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: In the adult, widely expressed with high
levels in heart, lung, kidney, skeletal muscle and testis.
CC -!- DEVELOPMENTAL STAGE: In the embryo, expression is first detected
on day 9 and increases up to day 18. Lower levels are found in
adult. At 9.5 dpc, expression is localised to the apical
ectodermal ridge (AER) of the developing fore- and hindlimb buds,
the telencephalon and the first brachial arch. At 10.5 dpc,
expression is also observed in the myotome and in sensory tissues
such as the nasal pit and optic vesicle.
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AB059617; BAB40968.1; -.
CC HSP; P00747; 1CEA.

```


[illegible]

or send an email to license@sib-sib.ch).

EMBL; X06290; CAA29618.1; -
 PIR; S00657; S00657.
 PDB; 1I71; 13-JUN-01.
 PDB; 1JFN; 28-JUN-02.
 PDB; 1KIV; 18-MAY-99.
 PDB; 3KIV; 18-MAY-99.
 PDB; 4KIV; 18-MAY-99.
 MEROPS; S01.226; -
 Genew; HGNC:6667; LPA.
 MIM; 152200; -
 GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.
 GO; GO:0008015; P:circulation; TAS.
 GO; GO:0009405; P:pathogenesis; TAS.
 InterPro; IPR001314; Chymotrypsin.
 InterPro; IPR000001; Kringle.
 InterPro; IPR001254; Ser:protease_Try.
 Pfam; PF00051; Kringle; 38.
 Pfam; PF00089; trypsin; 1.
 PRINTS; PR00722; CHYMOTRYPSIN.
 PRINTS; PR00018; KHINGLE.
 ProDom; PD000395; Kringle; 38.
 SMART; SM00130; KR; 38.
 SMART; SM00020; Tryp_Spc; 1.
 PROSITE; PS00021; KRINGLE_1; 38.
 PROSITE; PS00070; KRINGLE_2; 38.
 PROSITE; PS00240; TRYPSIN_DOM; 1.
 PROSITE; PS00134; TRYPSIN_HIS; 1.
 PROSITE; PS00135; TRYPSIN_SER; 1.
 Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
 Kringle; Repeat; Atherosclerosis; Signal; Polymorphism; 3D-structure.
 SIGNAL 1 19
 CHAIN 20 4548 APOLIPOPROTEIN(A).
 DOMAIN 20 130 KRINGLE TYPE IV, 1.
 DOMAIN 131 244 KRINGLE TYPE IV, 2.
 DOMAIN 245 358 KRINGLE TYPE IV, 3.
 DOMAIN 359 472 KRINGLE TYPE IV, 4.
 DOMAIN 473 586 KRINGLE TYPE IV, 5.
 DOMAIN 587 700 KRINGLE TYPE IV, 6.
 DOMAIN 701 814 KRINGLE TYPE IV, 7.
 DOMAIN 815 928 KRINGLE TYPE IV, 8.
 DOMAIN 929 1042 KRINGLE TYPE IV, 9.
 DOMAIN 1043 1156 KRINGLE TYPE IV, 10.
 DOMAIN 1157 1270 KRINGLE TYPE IV, 11.
 DOMAIN 1271 1384 KRINGLE TYPE IV, 12.
 DOMAIN 1385 1498 KRINGLE TYPE IV, 13.
 DOMAIN 1499 1612 KRINGLE TYPE IV, 14.
 DOMAIN 1613 1726 KRINGLE TYPE IV, 15.
 DOMAIN 1727 1840 KRINGLE TYPE IV, 16.
 DOMAIN 1841 1954 KRINGLE TYPE IV, 17.
 DOMAIN 1955 2068 KRINGLE TYPE IV, 18.
 DOMAIN 2069 2182 KRINGLE TYPE IV, 19.
 DOMAIN 2183 2296 KRINGLE TYPE IV, 20.
 DOMAIN 2297 2410 KRINGLE TYPE IV, 21.
 DOMAIN 2411 2524 KRINGLE TYPE IV, 22.
 DOMAIN 2525 2638 KRINGLE TYPE IV, 23.
 DOMAIN 2639 2752 KRINGLE TYPE IV, 24.
 DOMAIN 2753 2866 KRINGLE TYPE IV, 25.
 DOMAIN 2867 2980 KRINGLE TYPE IV, 26.
 DOMAIN 2981 3094 KRINGLE TYPE IV, 27.
 DOMAIN 3095 3208 KRINGLE TYPE IV, 28.
 DOMAIN 3209 3322 KRINGLE TYPE IV, 29.
 DOMAIN 3323 3436 KRINGLE TYPE IV, 30.
 DOMAIN 3437 3550 KRINGLE TYPE IV, 31.
 DOMAIN 3551 3664 KRINGLE TYPE IV, 32.
 DOMAIN 3665 3779 KRINGLE TYPE IV, 33.
 DOMAIN 3780 3894 KRINGLE TYPE IV, 34.
 DOMAIN 3895 3998 KRINGLE TYPE IV, 35.
 DOMAIN 3999 4112 KRINGLE TYPE IV, 36.
 DOMAIN 4113 4226 KRINGLE TYPE IV, 37.
 DOMAIN 4227 4340 KRINGLE TYPE V.
 SERINE PROTEASE.
 4341 4548

FT ACT_SITE 4369 4369 CHARGE RELAY SYSTEM.
 FT ACT_SITE 4412 4412 CHARGE RELAY SYSTEM.
 FT ACT_SITE 4498 4498 CHARGE RELAY SYSTEM.
 FT VARIANT 4193 4193 W->R (LOSS OF LYSINE-SEPHAROSE BINDING).
 FT /FTID=VAR_006633.
 SQ SEQUENCE 4548 AA; 501313 MW; 96921BE9A465C5F CRC64;
 Query Match 28.3%; Score 157; DB 1; Length 4548;
 Best Local Similarity 34.2%; Pred. No. 3.7e-09;
 Matches 40; Conservative 9; Mismatches 36; Indels 32; Gaps 6;
 QY 3 CYEGNGHFYRGKASTDTMGRCPLPWNSTVLTQTYHAHRSDAL---QLGLGHNYCRNPD 59
 DB 3896 CYRGDQSYRGTLSTITIGRTCSWSS-----MTPHWHRIPLYYPNAGLTR-NYCRNPD 3949
 QY 60 NRRRPWCYVQGLKPLV--QECMVHDC-----ADGKKPSPEE 96
 DB 3950 AEIRPMCYT---MDPSVRWEYCNLTCPVTSESVLTPTTVPVPSTEAPSEQAPPEK 4003
 RESULT 25
 KRML_HUMAN STANDARD; PRT; 475 AA.
 ID KRML_HUMAN Q96M08; Q9BY70; Q9UGS5; Q9UGU1;
 AC Q96M08; Q9BY70; Q9UGS5; Q9UGU1;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Kremen protein 1 precursor (Kringle-containing protein marking the eye
 DE and the nose) (Dickkopf receptor).
 GN KREMEN1 OR KREMEN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Nakamura T., Nakamura T.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
 RA Otsukumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
 RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
 RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
 RA Isogai T.;
 RT "NEDO human cDNA sequencing project";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
 RA Clump M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
 RA Burdill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 RA Dhama P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.,
 RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
 RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
 RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,
 RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
 RA Scott C.E., Senra H.K., Skuce C.D., Smalley S., Smith M.L.,
 RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,

FT	DOMAIN	416	475	CYTOLASMIC (POTENTIAL).
FT	DOMAIN	33	116	KRINGLE.
FT	DOMAIN	121	202	WSC.
FT	DOMAIN	216	323	CUB.
FT	CARBOHYD	47	47	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	61	61	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	219	219	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	295	295	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	335	335	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	347	347	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPPLIC	473	475	2).
FT				/FTID=VSP.003900.
FT	CONFLICT	29	30	MISSING (IN REF. 1).
FT	CONFLICT	206	206	I -> V (IN REF. 2).
FT	SEQUENCE	475 AA;	51898 MW;	B7E86FD80F96A04 CRC64;
ST				
ST	Query Match		28.2%;	Score 156; DB 1; Length 475;
ST	Best Local Similarity	44.1%;	Pred.No. 4.4e-10;	
ST	Matches 30;	Conservative	8;	Mismatches 26; Indels 4; Gaps 2;
QY	3	CYEGNGHGFYRGKASTDTM--GRPCLPWNSATVLOQTYHAHRSDALQLGLGHXNYCRPNPDN	60	
Db	34	CFTANGADYEGTQNTWALQGGKPCLFMNE--TFQHPYNTLKPNGEGGLGEHNYCRPNPDG	91	
QY	61	RRRPKCYV 58		
Db	92	DVSPWCYV 99		
ST				
ST	RESULT 26			
ST	FA12 BOVIN			
ID	FA12 BOVIN	STANDARD;	PRT;	593 AA.
AC	P98140;			
DT	01-FEB-1996	(Rel. 33, Created)		
DT	01-FEB-1996	(Rel. 33, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)			
DE	(HAF) (Fragment).			
GN	F12.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RC	MEDLINE=94242782; PubMed=8186251;			
RA	Shibuya Y., Samba U., Okabe H., Kambara T., Yamamoto T.;			
RT	"Primary structure of bovine Hageman factor (blood coagulation factor			
RT	XIII): comparison with human and guinea pig molecules."			
RL	Biochim. Biophys. Acta 1206:63-70(1994).			
RN	[2]			
RP	SEQUENCE OF 10-21; 350-364 AND 525-550.			
RP	MEDLINE=77182112; PubMed=861210;			
RA	Fujikawa K., Walsh A.K., Davie W.E.;			
RT	"Isolation and characterization of bovine factor XII (Hageman			
RT	factor)."			
RL	Biochemistry 16:2270-2278(1977).			
CC	-1- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN			
CC	THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE			
CC	GENERATION OF BRADYKININ AND ANGIOTENSIN.			
CC	-1- CATALYTIC ACTIVITY: Cleaves selectively Arg--ile bonds in factor			
CC	XII to form factor VIIa and factor XI to form factor Xla.			
CC	-1- PTM: O- AND N-GLYCOSYLATED (BY SIMILARITY)			
CC	-1- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A			
CC	COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY			
CC	FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST			
CC	TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR			
CC	XIIA ACTIVATES FACTOR XI TO FACTOR XIA. BOVINE FACTOR XI IS			
CC	CLEAVED ONLY TO ALPHA-FACTOR XIIA AS IT LACKS THE TRYPSIN/			
CC	KALLIKREIN CLEAVAGE SITE.			

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: Contains 2 EGF-like domains.
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: S70164; AAB30804.2; -;
 DR PIR: S45281; S45281.
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.211; -;
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR006209; EGF-like.
 DR InterPro: IPR000083; Fibnctnl.
 DR InterPro: IPR000562; FN Type II.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00008; EGF; 2.
 DR Pfam: PF00039; fn1; 1.
 DR Pfam: PF00040; fn2; 1.
 DR Pfam: PF00051; kringle; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00013; FNTYPEII.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000995; FN Type II; 1.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART: SM00181; EGF; 2.
 DR SMART: SM00058; FN1; 1.
 DR SMART: SM00059; FN2; 1.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PS00022; EGF 1; 2.
 DR PROSITE: PS01186; EGF 2; FALSE NEG.
 DR PROSITE: PS01253; FIBRONECTIN I; 1.
 DR PROSITE: PS00023; FIBRONECTIN 2; 1.
 DR PROSITE: PS00021; KRINGLE 1; 1.
 DR PROSITE: PS00070; KRINGLE 2; 1.
 DR PROSITE: PS0240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR Glycoprotein: Blood coagulation; Plasma; Serine protease;
 KW Hydrolyase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
 FT NON_TER 1
 FT SIGNAL <1 9
 FT CHAIN 10 349
 FT CHAIN 350 593
 FT DOMAIN 37 78
 FT DOMAIN 84 121
 FT DOMAIN 123 163
 FT DOMAIN 164 200
 FT DOMAIN 207 287
 FT DOMAIN 297 333
 FT DOMAIN 337 389
 FT DOMAIN 389 438
 FT ACT_SITE 438 438
 FT ACT_SITE 541 541
 FT DISULFID 88 100
 FT DISULFID 94 109
 FT DISULFID 111 120
 FT DISULFID 125 153
 FT DISULFID 151 160
 FT DISULFID 168 179
 FT DISULFID 173 188
 FT DISULFID 190 199

FT DISULFID 207 287 BY SIMILARITY.
 FT DISULFID 230 269 BY SIMILARITY.
 FT DISULFID 258 282 BY SIMILARITY.
 FT DISULFID 336 463 BY SIMILARITY.
 FT DISULFID 374 390 BY SIMILARITY.
 FT DISULFID 382 452 BY SIMILARITY.
 FT DISULFID 413 416 BY SIMILARITY.
 FT DISULFID 479 547 BY SIMILARITY.
 FT DISULFID 510 526 BY SIMILARITY.
 FT DISULFID 537 568 BY SIMILARITY.
 FT CARBOHYD 99 O-LINKED (FUC) (BY SIMILARITY).
 FT CARBOHYD 241 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 263 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 410 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 593 AA; 65148 MW; 721592BA792BD61F CRC64;
 Query Match 28.1%; Score 155.5; DB 1; Length 593;
 Best Local Similarity 38.4%; Pred. No. 6.3e-10;
 Matches 33; Conservative 9; Mismatches 37; Indels 7; Gaps 3;
 QY 2 TCYE--GNGHEFYRGKASTDTMTGRPCLPWNSATVLQOTY-HAHRSDALQLGLGKHNYCRNP 58
 DB 206 SCYDDRDRLSYRGVAGTTLSCAPQSWAS----EATYWNVTAEQVLNWLGLDGHAFCRNP 261
 QY 59 DNRRRPWCYVQVGLKPLVQECMVHDC 84
 DB 262 DNDTRPWCFFWKDRLSNWYCYRLAPC 287
 RESULT 27
 PLMN_BOVIN
 ID PLMN_BOVIN STANDARD; PRT; 812 AA.
 AC P06868; Q28162;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasmimogen precursor (EC 3.4.21.7).
 GN PLG.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Berglund L., Andersen M.D., Petersen T.E.;
 RT "Cloning and characterization of the bovine plasminogen cDNA.";
 RL Int. Dairy J. 5:593-603(1995).
 RN [2]
 RP SEQUENCE OF 27-812, AND CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=85203906; PubMed=3846532;
 RA Schaller J., Moser P.W., Danneberger-Muller G.A.K., Rosselet S.J.,
 RT Kampfer U., Rickli E.E.;
 RL "Complete amino acid sequence of bovine plasminogen. Comparison with
 human plasminogen.";
 RL Eur. J. Biochem. 149:267-278(1985).
 RN [3]
 RP SEQUENCE OF 706-812 FROM N.A.
 RX MEDLINE=85023311; PubMed=6148961;
 RA Malinowski D.P., Sadler J.E., Davie E.W.;
 RT "Characterization of a complementary deoxyribonucleic acid coding for
 human and bovine plasminogen.";
 RL Biochemistry 23:4243-4250(1984).
 RN [4]
 RP CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=88185329; PubMed=3356193;
 RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
 RA Gerwig G.J., van Halbeek H., Vliegthart J.F.;
 RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
 plasminogen. Species specificity in relation to sialylation and
 fucosylation patterns.";
 RL Eur. J. Biochem. 173:57-63(1988).

CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
 CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
 CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
 CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
 CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
 CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
 CC LAMININ AND VON WILLEBRAND FACTOR.
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 CC higher selectivity than trypsin. Converts fibrin into soluble
 CC products.
 CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
 CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
 CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
 CC -!- PTM: N-LINKED GLYCAN CONTAIN N-ACETYLACTOSAMINE AND SIALIC ACID.
 CC O-LINKED GLYCANS CONSIST OF GAL-GALNAc RESIDUES DISACCHARIDE WITH IS
 CC MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES (MICROHETEROGENEITY).
 CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
 CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -!- SIMILARITY: Contains 5 kringle domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X79402; CAA55939.1; -;
 CC EMBL; K02935; AAA30714.1; -;
 CC PIR; S45046; PLEB.
 CC HSSP; P00747; 2PK4.
 CC MEROPS; S01.233; -;
 CC GlycoSuiteDB; P06868; -;
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR003014; PAN.
 CC InterPro; IPR003609; PAN_app.
 CC InterPro; IPR003966; Prothrombin.
 CC InterPro; IPR001254; Ser.protease_Try.
 CC Pfam; PF00051; Kringle; 5.
 CC Pfam; PF00024; PAN; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC PRINTS; PR01505; PROTHROMBIN.
 CC ProDom; PD000395; Kringle; 5.
 CC SMART; SM00130; KR; 5.
 CC SMART; SM00473; PAN AP; 1.
 CC SMART; SM00020; Tryp_SPC; 1.
 CC PROSITE; PS00021; KRINGLE_1; 5.
 CC PROSITE; PS00070; KRINGLE_2; 5.
 CC PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 CC Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
 CC Signal.
 CC -----
 CC 1 26
 CC CHAIN 27 812 PLASMINOGEN.
 CC CHAIN 27 583 PLASMIN HEAVY CHAIN A.
 CC CHAIN 584 812 PLASMIN LIGHT CHAIN B.
 CC DOMAIN 110 188 KRINGLE 1.
 CC DOMAIN 192 269 KRINGLE 2.
 CC DOMAIN 282 359 KRINGLE 3.
 CC DOMAIN 384 461 KRINGLE 4.
 CC DOMAIN 485 564 KRINGLE 5.
 CC DOMAIN 584 812 SERINE PROTEASE.
 CC CARBOHYD 315 315 N-LINKED (GLCNAC. . .).
 CC -----
 CC CARBOHYD 365 365 O-LINKED (GALNAC. . .).
 CC -----

FT ACT_SITE 624 624 /FTid-CAR_000015.
 FT ACT_SITE 667 667 CHARGE RELAY SYSTEM.
 FT ACT_SITE 762 762 CHARGE RELAY SYSTEM.
 FT CONFLICT 335 335 N -> D (IN REF. 2).
 FT CONFLICT 516 516 Q -> H (IN REF. 2).
 FT CONFLICT 555 555 P -> L (IN REF. 2).
 FT CONFLICT 744 744 T -> R (IN REF. 3).
 SQ SEQUENCE 812 AA; 91216 MW; 38AA6A691E220946 CRC64;
 Query Match 27.4%; Score 152; DB 1; Length 812;
 Best Local Similarity 37.8%; Pred. No. 2.2e-09;
 Matches 37; Conservative 11; Mismatches 34; Indels 16; Gaps 6;
 QY 3 CYEGNGHYRGRKASTDTMGRPCLPWNSATVLOQTYHAH---RSDALQLGLGHKHYCRNPD 59
 DB 384 CYHNGQSYRGTSSTITGRKQSWSS-----MTPRHLLKTPENYPNAGL-TMNYCRNPD 437
 QY 60 NRRRWCVQVGLKPLV--QECMVHDCADGKK--PSSP 93
 DB 438 ADKSPACVT---TDRVWFECNLKCKSETPEQVPAAP 472
 RESULT 28
 PLMN RAT
 ID PLMN RAT STANDARD; PRT; 169 AA.
 AC Q01177;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasminogen (EC 3.4.21.7) (Fragment).
 GN PLG.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=91250378; PubMed=1645711;
 RA Kanalas J.J., Makker S.P.;
 RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
 RT receptor site for plasminogen.";
 RL J. Biol. Chem. 266:10825-10829(1991).
 CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
 CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
 CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
 CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
 CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
 CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
 CC LAMININ AND VON WILLEBRAND FACTOR.
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 CC higher selectivity than trypsin. Converts fibrin into soluble
 CC products.
 CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
 CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
 CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
 CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
 CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -!- SIMILARITY: Contains 5 kringle domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M62832; AAA41884.1; -;
 CC PIR; A40522; A40522.


```
DR HSP; P00747; LPMK.
DR MEROPS; S01.233; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; P00051; kringle; 2.
DR PRINTS; P00018; KRINGLE.
DR PRINTS; P00018; KRINGLE.
DR ProDom; P0000395; Kringle; 2.
DR SMART; SM00130; KR; 1.
DR PROSITE; P500021; KRINGLE 1; 1.
DR PROSITE; P500070; KRINGLE 2; 2.
DR PROSITE; P500240; TRYPSIN_DOM; PARTIAL.
DR PROSITE; P500134; TRYPSIN_HIS; PARTIAL.
DR PROSITE; P500135; TRYPSIN_SER; PARTIAL.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Repeat.
FT NON_TER 1 1
FT DOMAIN <1 10 KRINGLE 3 (BY SIMILARITY).
FT DOMAIN 34 112 KRINGLE 4 (BY SIMILARITY).
FT DOMAIN 139 >169 KRINGLE 5 (BY SIMILARITY).
FT DISULFID 34 112 BY SIMILARITY.
FT DISULFID 55 95 BY SIMILARITY.
FT DISULFID 83 107 BY SIMILARITY.
FT NON_TER 169 169
SQ SEQUENCE 169 AA; 18401 MW; 77A54214C49D010C CRC64;

Query Match 27.1%; Score 150; DB 1; Length 169;
Best Local Similarity 34.3%; Pred. No. 7e-10;
Matches 37; Conservative 14; Mismatches 31; Indels 26; Gaps 7;

QY 3 CYEGNGHYVGRKASTDMGRPCPLPWNASTVLQOTVHAHRSDALQL---GLGKHNYCRNPD 59
Dd 34 CYGNGKSVGRGTSITNTGKCSQW-----VSMTPSHSKTKTANFPDSDL-ENWYCRNPD 87
QY 60 N-RRRPWCYVQGLKPLV--QECMHDCAD-----GKKPSSP 93
Dd 88 NDQRGWCFT---TDPVSRWEYCNLKRCSGTGGVAESAIVPQVPSAP 132

RESULT 29
PLWN_MACMU STANDARD; PRT; 810 AA.
AC P12545;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7).
GN PLG.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89174660; PubMed=2925643;
RA Tomlinson J.E., McLean J.W., Lawn R.M.;
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
RT synthesis."
RL J. Biol. Chem. 264:5957-5965 (1989).
CC -1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
```

```

FT DISULFID 152 176 BY SIMILARITY.
FT DISULFID 185 262 BY SIMILARITY.
FT DISULFID 188 316 BY SIMILARITY.
FT DISULFID 206 245 BY SIMILARITY.
FT DISULFID 234 257 BY SIMILARITY.
FT DISULFID 275 352 BY SIMILARITY.
FT DISULFID 296 335 BY SIMILARITY.
FT DISULFID 324 347 BY SIMILARITY.
FT DISULFID 377 454 BY SIMILARITY.
FT DISULFID 398 437 BY SIMILARITY.
FT DISULFID 426 449 BY SIMILARITY.
FT DISULFID 481 560 BY SIMILARITY.
FT DISULFID 502 543 BY SIMILARITY.
FT DISULFID 531 555 BY SIMILARITY.
FT DISULFID 567 685 BY SIMILARITY.
FT DISULFID 577 585 BY SIMILARITY.
FT DISULFID 607 623 BY SIMILARITY.
FT DISULFID 699 766 BY SIMILARITY.
FT DISULFID 729 745 BY SIMILARITY.
FT DISULFID 756 784 BY SIMILARITY.
FT CARBOHYD 365 365 O-LINKED (GALNAc... ) (BY SIMILARITY).
SQ SEQUENCE 810 AA; 90255 MW; A75EC51A1A0F24A CRC64;

Query Match 26.9%; Score 149; DB 1; Length 810;
Best Local Similarity 36.4%; Pred. No. 4.7e-09;
Matches 36; Conservative 12; Mismatches 35; Indels 16; Gaps 6;

QY 3 CYEGNGHYRGKASTDTWGRCLPNSATVLOQTYHAHR---SPALQLGLGKHNYCENPD 59
DB 377 CYHGDGQSGYRGTSSTTTGKRCQSWSS-----MTPHHEKTPENFNPAGL-TNYICRNPD 430
QY 60 NRRREWCYVGVGLKPLV--QECMVHDC--DGKRPSPSP 94
DB 431 ADKGPWCFT---TDFSVRWEYCNLKCSGTGCSVAAPP 466

RESULT 30
KRM2 HUMAN
ID KRM2 HUMAN STANDARD; PRT; 462 AA.
AC Q8NCW0; Q8NCW1; Q96GLS; Q9BTP9;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 2 precursor (Kringles-containing protein marking the eye
DE and the nose) (Dickkopf receptor 2).
GN KRMEN2 OR KRM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RA Tanaka S., Sugimachi K.;
RT "Human Kremen2 and Wnt signaling.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]_
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RC TISSUE=Ovarian carcinoma;
RA Isozai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hiko Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Masuno Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]_
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RC TISSUE=Brain, and Uterus;
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

```

```

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling. Forms a ternary complex with
CC Dkk1 and LRP6 and induces rapid endocytosis and removal of the Wnt
CC receptor LRP6 from the plasma membrane (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1;
CC IsoId=Q8NCW0-1; Sequence=Displayed;
CC Name=2; Synonyms=Kremen2a;
CC IsoId=Q8NCW0-2; Sequence=VSP_050509, VSP_050510;
CC Name=3; Synonyms=Kremen2b;
CC IsoId=Q8NCW0-3; Sequence=VSP_050511, VSP_050512;
CC Name=4; Synonyms=Kremen2c;
CC IsoId=Q8NCW0-4; Sequence=VSP_050513, VSP_050514;
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 Kringles domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB086405; BAC00872.1; -
DR EMBL; AB086355; BAC00823.1; -
DR EMBL; AB086356; BAC00824.1; -
DR EMBL; AB086357; BAC00825.1; -
DR EMBL; AK027669; BAB5281.1; -
DR EMBL; AK075033; BAC11365.1; -
DR EMBL; BC003533; AAH03533.1; -
DR EMBL; BC009383; AAH09383.1; -
DR HSSP; P00750; 1PK2.
DR Genew; HGNC:18797; KRMEN2.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR000001; Kringles.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; Kringles; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE5.
DR Prodom; PD000395; Kringles; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00321; WSC; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE-2; 1.
KW Wnt signaling pathway; Glycoprotein; Kringles; Signal; Transmembrane;
KW Alternative splicing.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 462 KRMEN PROTEIN 2.
FT DOMAIN 26 364 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 365 387 POTENTIAL.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:23 ; Search time 11.8119 Seconds
(without alignments)
2097.294 Million cell updates/sec

Title: US-09-880-503-9

Perfect score: 554

Sequence: 1 KTCYEGNGHYRGKASTDTM.....QECMVHDCADGKKSPPEE 96

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	495	89.4	154	4 Q96SE8	Q96se8 homo sapien
2	440	79.4	433	6 Q8MIL0	Q8mil0 oryctolagus
3	440	79.4	433	6 Q8MIL0	Q8mil0 oryctolagus
4	422	76.2	157	6 Q9TVA8	Q9tva8 bos taurus
5	408	73.6	231	11 Q8C6L2	Q8c6l2 mus musculus
6	337.5	60.9	214	6 Q9XT70	Q9xt70 oryctolagus
7	236.5	42.7	564	6 Q8MKB1	Q8mbk1 oryctolagus
8	226	40.8	516	4 Q9BU99	Q9bu99 homo sapien
9	219	39.5	559	11 Q91VP2	Q91vp2 mus musculus
10	218.5	39.4	90	4 Q8NG20	Q8ng20 homo sapien
11	218.5	39.4	395	4 Q9BZW1	Q9bzw1 homo sapien
12	211	38.1	562	6 Q8SQ23	Q8sq23 sus scrofa
13	210	37.9	517	11 Q8K0D2	Q8kd02 mus musculus
14	199.5	36.0	560	4 Q14520	Q14520 homo sapien
15	191	34.5	202	13 Q90675	Q90675 gallus galli
16	188	33.9	653	11 Q8VCS4	Q8vcs4 mus musculus

17	187.5	33.8	597	11	O35727	O35727 mus musculus
18	178.5	32.2	616	6	O97507	O97507 sus scrofa
19	172.5	31.1	421	13	Q8AXX3	Q8axx3 xenopus lae
20	170.5	30.8	615	4	Q8IZZ5	Q8izz5 homo sapien
21	167	30.1	947	13	Q8AXY6	Q8axy6 gallus galli
22	158	28.5	452	13	Q90Y90	Q90y90 xenopus lae
23	158	28.5	454	6	O46506	O46506 papio hamad
24	157	28.3	113	4	Q9UIR5	Q9uir5 homo sapien
25	154	27.8	806	6	O18783	O18783 macropus eu
26	153.5	27.7	801	11	Q8K0Q8	Q8k0q8 mus musculus
27	153.5	27.7	944	11	Q8C3W2	Q8c3w2 mus musculus
28	153.5	27.7	944	11	Q8BSF6	Q8bsf6 mus musculus
29	153	27.6	113	4	Q9UIR7	Q9uir7 homo sapien
30	151.5	27.3	359	6	Q8WMR1	Q8wmr1 canis famil
31	150	27.1	812	11	Q9ROW3	Q9row3 rattus norv
32	149	26.9	716	13	Q91691	Q91691 xenopus lae
33	148.5	26.8	381	4	Q8N2J4	Q8n2j4 homo sapien
34	148.5	26.8	399	4	Q96GJ8	Q96gl8 homo sapien
35	148.5	26.8	420	4	Q9BTB9	Q9btp9 homo sapien
36	148.5	26.8	424	4	Q8NCW1	Q8ncw1 homo sapien
37	148.5	26.8	462	4	Q8NCW0	Q8ncw0 homo sapien
38	148	26.7	105	4	Q9UIR8	Q9uir8 homo sapien
39	147.5	26.6	567	4	Q13208	Q13208 homo sapien
40	147.5	26.6	648	4	Q9H1V4	Q9hlv4 homo sapien
41	147.5	26.6	716	11	Q9IXG8	Q9ixg8 mus musculus
42	147	26.5	113	4	Q9UIR6	Q9uir6 homo sapien
43	146.5	26.4	461	11	Q8K1S7	Q8kl57 mus musculus
44	146.5	26.4	810	4	Q15146	Q15146 homo sapien
45	146	26.4	393	4	Q9BRB6	Q9brb6 homo sapien

ALIGNMENTS

RESULT 1

Q96SE8 PRELIMINARY, PRT; 154 AA.
AC Q96SE8
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator amino-terminal fragment.
GN ATF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu J., Bai X., Ruan C.;
RT "Cloning and expression of the amino-terminal fragment of human urokinase-type plasminogen activator."
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
RT "Overexpression of the amino-terminal fragment of human urokinase-type plasminogen activator in breast cancer cells results in decreased tumor invasion, growth and angiogenesis.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY029537; AAK38734.1; -.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.

```
KW Glycoprotein; Kinase; Kringle.
SQ SEQUENCE 154 AA; 17305 MW; A3CCF2FCFF505572 CRC64;

Query Match 89.4%; Score 495; DB 4; Length 154;
Best Local Similarity 98.9%; Pred. No. 2.4e-51;
Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFGKASTDTWGRCLPNSATVLOQTYHAHRSDALQLGLGKHNCRPN 60
Dy 68 KTCYEGNGHFGKASTDTWGRCLPNSATVLOQTYHAHRSDALQLGLGKHNCRPN 127

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87
Dy 128 RRRPWCYVQVGLKLLVQECMVHDCADG 154

RESULT 2
QBMIL0 PRELIMINARY; PRT; 433 AA.
AC QBMIL0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN PLAU.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22155943; PubMed=12149463;
RA Falkenberg M., Tom C., Deyoung M.B., Wen S., Linnemann R.,
RA Dichek D.A.;
RT "Increased expression of urokinase during atherosclerotic lesion
RT development causes arterial constriction and lumen loss, and
RT accelerates lesion growth";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AY122285; AAM83187.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Serine protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00089; trypsin; 1.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65E64F3641549B0 CRC64;

Query Match 79.4%; Score 440; DB 6; Length 433;
Best Local Similarity 78.7%; Pred. No. 2.9e-44;
Matches 74; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFGKASTDTWGRCLPNSATVLOQTYHAHRSDALQLGLGKHNCRPN 60
Dy 70 KTCYEGNGHFGKASTDTWGRCLPNSATVLOQTYHAHRSDALQLGLGKHNCRPN 129

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPP 94
Dy 130 QRRPWCYVQVGLKQLIQECKVHDCSSGKKPALPP 163

RESULT 3
QBMHY7 PRELIMINARY; PRT; 433 AA.
AC QBMHY7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
RT complete cds.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AY029517; AAK40239.1; -.
DR EMBL; AB087224; BAC02685.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Serine protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 6DD35A371010A6EE CRC64;

Query Match 79.4%; Score 440; DB 6; Length 433;
Best Local Similarity 78.7%; Pred. No. 2.9e-44;
Matches 74; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFGKASTDTWGRCLPNSATVLOQTYHAHRSDALQLGLGKHNCRPN 60
Dy 70 KTCYEGNGHFGKASTDTWGRCLPNSATVLOQTYHAHRSDALQLGLGKHNCRPN 129

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPP 94
Dy 130 QRRPWCYVQVGLKQLIQECKVHDCSSGKKPALPP 163

RESULT 4
Q9TVAS PRELIMINARY; PRT; 157 AA.
AC Q9TVAS;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
```

```

OX NCBI_TaxID=9913;
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=21071388; PubMed=11204721;
RT Balcerzak D., Quereguesser L., Dixon W.T., Baracos V.E.;
RA "Coordinate expression of matrix-degrading proteinases and their
RT activators and inhibitors in bovine skeletal muscle.";
RL J. Anim. Sci. 79:94-107(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF144761; AAD30301.1; -.
DR HSSP; P00749; LURK.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
KW Glycoprotein; Kinase; Kringle.
FT NON_TER 1
FT NON_TER 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7 CRC64;

Query Match 76.2%; Score 422; DB 6; Length 157;
Best Local Similarity 75.0%; Pred. No. 1.3e-42;
Matches 72; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCPWPNSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60
DB 36 KTCYCGNGSHYRGKANRDLSSGPCLAWSPTVLLKMYHAHRSDALQLGLGKHNCRNPDN 95
QY 61 RRRPCYVQVGLKPLVQECNVHDCADGKXPSPPE 96
DB 96 QRRPCYVQVGLKPLVQECNVHDCADGKXPSPPREK 131

RESULT 5
Q8C6L2 PRELIMINARY; PRT; 231 AA.
AC Q8C6L2;
DT 01-WAR-2003 (Tremblrel. 23, Created)
DT 01-WAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-WAR-2003 (Tremblrel. 23, Last annotation update)
DE Plasmogen activator (fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10990;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK054349; BAC35743.1; -.
FT NON_TER 231
SQ SEQUENCE 231 AA; 25510 MW; 25E8980A692737F2 CRC64;

Query Match 73.6%; Score 408; DB 11; Length 231;
Best Local Similarity 70.8%; Pred. No. 9.6e-41;
Matches 68; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCPWPNSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60
DB 69 KTCYHGNGSHYRGKANTDKRGPCPLAWNAPAVLQFPYNAHRPDAISLGLGKHNCRNPDN 128
QY 61 RRRPCYVQVGLKPLVQECNVHDCADGKXPSPPE 96
```

```

Db 129 OKRPWCYVQVGLKPLVQECNVHDCSLSKPSSSDQ 164
:::|||||:::|||||:::|||||:::|||||:::|||||:::
Q9XT70 PRELIMINARY; PRT; 214 AA.
AC Q9XT70;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator (fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Yin J., Idell S.;
RT "Partial mRNA of rabbit uPA.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097647; AAD39351.1; -.
DR HSSP; P00749; 1EJN.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00089; trypsin; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
FT NON_TER 214
SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

Query Match 60.9%; Score 337.5; DB 6; Length 214;
Best Local Similarity 78.7%; Pred. No. 2.4e-32;
Matches 59; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

QY 20 MGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLKPLVQEC 79
DB 1 MDRPCLAWSANVLTKYHAHRPDALQLGLGKHNCRNPNRPHQRPWCYVQVGLKPLVQEC 60
:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 80 MVHDCADGKXPSPPE 94
:::|||||:::
DB 61 KVHD-SSGKKPALPP 74

RESULT 7
Q8MKBI PRELIMINARY; PRT; 564 AA.
AC Q8MKBI;
ID Q8MKBI;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Tissue-type plasminogen activator.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
```

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 CC EMBL; AY029518; AAK40240.1; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR000083; Fibinctnl.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF_1.
 DR Pfam; PF00039; fnl; 1.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 2.
 DR PROSITE; PS50070; KRINGLE_2; 2.
 DR PROSITE; PS50240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KW Serine protease.
 SQ SEQUENCE 564 AA; 62726 MW; 459D8BAC6D4A937C CRC64;

Query Match 42.7%; Score 236.5; DB 6; Length 564;
 Best Local Similarity 46.8%; Pred. No. 8.8e-20;
 Matches 44; Conservative 12; Mismatches 37; Indels 1; Gaps 1;

QY 3 CYEGNGHFYRGKASTDTMGPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDR 62
 DB 217 CYLGKQAYRGTHSLTSGASCLPWNMLLVGSKYARQSNAAELGLGKHNYCRNPDR 276
 QY 63 RPKWYQVGLKPLVQECMVHDCAD-GKKPSRPE 95
 DB 277 KPWCHVLKNEKLVIEYCDVPCQATCGLRQDKQKP 310

RESULT 8
 Q9BU99 PRELIMINARY; PRT; 516 AA.
 AC Q9BU99
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to plasminogen activator, tissue.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 DR EMBL; BC002795; AAH02795.1; -.
 DR HSSP; PF00750; IASH.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF_1.
 DR Pfam; PF00051; kringle; 2.

DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 2.
 DR PROSITE; PS50070; KRINGLE_2; 2.
 DR PROSITE; PS50240; TRYPsin_DOM; 1.
 DR PROSITE; PS00134; TRYPsin_HIS; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KW Serine protease.
 SQ SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;

Query Match 40.8%; Score 226; DB 4; Length 516;
 Best Local Similarity 47.7%; Pred. No. 1.4e-18;
 Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDR 61
 DB 80 TCYEDQGISYRGTWSTABSGAECTWNSALLAQKPYSGRRPDALRLGLGNHNYCRNPDR 139
 QY 62 RPKWYQVGLKPLVQECMVHDCADG 87
 DB 140 SKPCYVFPKAGKYSSEFCSTPACSEG 165

RESULT 9
 Q91VP2 PRELIMINARY; PRT; 559 AA.
 ID Q91VP2
 AC Q91VP2
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to plasminogen activator, tissue.
 GN PLAT.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast tumor;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 DR EMBL; BC011256; AAH11256.1; -.
 DR HSSP; PF00761; LAN1.
 DR MGD; MGI:97610; Plat.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR000083; Fibinctnl.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF_1.
 DR Pfam; PF00039; fnl; 1.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.

```
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 559 AA; 63122 MW; 8CCBE2BDB94514D9 CRC64;

Query Match 39.5%; Score 219; DB 11; Length 559;
Best Local Similarity 46.0%; Pred. No. 1.1e-17;
Matches 40; Conservative 11; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCVEGNHFRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 61
DB 123 TCVEGQITVGGTSTAESGAECINMNSVLSKPYNARRPNAIKLGLGNHNYCRNPDRD 182
QY 62 RRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 183 LKPCYVFKAGKYTTERCFSPACKPKG 209

RESULT 10
Q8NG20 PRELIMINARY; PRT; 90 AA.
AC Q8NG20; 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DE Plasminogen/activator kringle.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RA Dou D.;
RT "Production of kringle fragment.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF282882; AAM52248.1;
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
KW Glycoprotein; Kringle.
SQ SEQUENCE 90 AA; 5804 MW; A33887F9DF4C7B1 CRC64;

Query Match 39.4%; Score 218.5; DB 4; Length 90;
Best Local Similarity 49.4%; Pred. No. 1.6e-18;
Matches 41; Conservative 5; Mismatches 36; Indels 1; Gaps 1;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 62
DB 8 CYFGNGSAYRGTHSLTESGASCLPWNMILIGKYVTAQNPSAQLGLGKHNYCRNP 67
QY 63 RPPWCYVQVGLKPLVQECMVHDC 85
DB 68 KPCYV-TNPKLYDCVPCQA 89

RESULT 11
Q9B2W1 PRELIMINARY; PRT; 395 AA.
AC Q9B2W1; 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Neonatal thrombolytic agent alpha-form (fragment).
```

```
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RA Dou D.;
RT "A brain-type plasminogen activator.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF260825; AAK11956.1;
DR HSSP; P00750; 1PK2.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON_TER 395
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 39.4%; Score 218.5; DB 4; Length 395;
Best Local Similarity 44.7%; Pred. No. 8.5e-18;
Matches 42; Conservative 10; Mismatches 41; Indels 1; Gaps 1;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 62
DB 48 CYFGNGSAYRGTHSLTESGASCLPWNMILIGKYVTAQNPSAQLGLGKHNYCRNP 107
QY 63 RPPWCYVQVGLKPLVQECMVHDCAD-GKKPSSPPE 95
DB 108 KPCWCHLVKNRRLTWEXCDVPSCGGLROYSQPQ 141

RESULT 12
Q8SQ23 PRELIMINARY; PRT; 562 AA.
AC Q8SQ23; 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RA Dou D.;
RT "Tissue=Enamel organ;
RT Ding Y., Xue J., Bartlett J.D.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF364605; AAM00297.1;
DR HSSP; P00761; 1ANI.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR006210; IEGF.
```


100

```

DR      PRODOM; PD0000395; Kringie; 1.
DR      PROSITE: PS00022; EGF 1; 3

```

100

```
DR PROSITE; PS00186; EGF_2; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 560 AA; 62671 MW; 5C1907230784ACD4 CRC64;

Query Match 36.0%; Score 199.5; DB 4; Length 560;
Best Local Similarity 40.2%; Pred. No. 2.4e-15;
Matches 39; Conservative 15; Mismatches 40; Indels 3; Gaps 2;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPMNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDRR 62
DB 194 CYVGDGYSYRGKRNRTVNHQACLYWNSHLLQYNFMFEDAEETHGIGEHNFCRNPDADE 253

QY 63 RPMCYYQVGLKPLVQE-CMVHDC--DGKKSPSPPEE 96
DB 254 KPCWFIKVTNDKWKVEYCDVSACSQDVAYPEESPTE 290

RESULT 15
Q90675 PRELIMINARY; PRT; 202 AA.
AC Q90675;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tissue-type plasminogen activator (fragment).
GN TPA.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=White Leghorn;
RP Johnson A.L., Bridham J.T., Anthony R.V.;
RX MEDLINE=97199025; PubMed=9047000;
RT "Expression of avian urokinase and tissue-type plasminogen activator
messenger ribonucleic acid during follicle development and atresia.";
RL Biol. Reprod. 56:581-588(1997).
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; U31988; AAA74955.1; -.
DR HSP; P00750; IRTF.
DR MEROPS; S01.232; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 1.
DR SMART; SM00130; KR; 1.
DR HSP; P00750; IRTF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTPEPI.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 202
SQ SEQUENCE 202 AA; 23148 MW; 049DB42941D83AFC CRC64;

Query Match 34.5%; Score 191; DB 13; Length 202;
Best Local Similarity 43.9%; Pred. No. 7.8e-15;
Matches 36; Conservative 7; Mismatches 39; Indels 0; Gaps 0;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPMNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDRR 62
DB 41 CYTGNGLAYRGTRTSYKSGSCLPWFVLTSLIYTALEQRALGLGKHNCRNPDDGA 100
```

```
QY 63 RPMCYYQVGLKPLVQECMVHDC 84
DB 101 QPCHVHWKDRQLTWYCDVPOC 122

RESULT 16
Q8VCS4 PRELIMINARY; PRT; 653 AA.
AC Q8VCS4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 70.6 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAI19376.1; -.
DR HSP; P00761; IAN1.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000893; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTPEPI.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match 33.9%; Score 188; DB 11; Length 653;
Best Local Similarity 33.0%; Pred. No. 6.6e-14;
Matches 35; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPMNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDRR 62
DB 283 CFLGNGTEYRGVASTAASGLSCLAWNSDLLYQELHVDVAAAVLLGLGPHAYCRNPKDE 342

QY 63 RPMCYY 68
DB 343 RPMCYY 348
```

RESULT 17

O35727 PRELIMINARY; PRT; 597 AA.

AC O35727;

DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Factor XII.

GN F12.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA Schloesser M., Schwager S., Engel W.;

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; X99571; CAA67891.1; -

DR HSSP; P00760; IAQ7.

DR MEROPS; S01-211; -

DR MGD; MGI:1891012; F12.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR006209; EGF-like.

DR InterPro; IPR000083; Fibrinctn1.

DR InterPro; IPR000562; FN Type II.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00039; fn1; 1.

DR Pfam; PF00040; fn2; 1.

DR Pfam; PF00051; kringle; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00013; ENTPEII.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000995; FN Type II; 1.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00181; EGF; 2.

DR SMART; SM00058; FN1; 1.

DR SMART; SM00059; FN2; 1.

DR SMART; SM00130; KR; 1.

DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS00022; EGF_1; 2.

DR PROSITE; PS01186; EGF_2; 1.

DR PROSITE; PS01253; FIBRONECTIN_1; 1.

DR PROSITE; PS00023; FIBRONECTIN_2; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00021; KRINGLE_2; 1.

DR PROSITE; PS00070; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

DR PROSITE; PS00240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;

KW Serine protease.

SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0PBA CRC64;

Query Match 33.8%; Score 187.5; DB 11; Length 597;

Best Local Similarity 37.7%; Pred. No. 6.9e-14;

Matches 40; Conservative 10; Mismatches 39; Indels 17; Gaps 3;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLCQTY-HAHRSDALQLGLGKHNYPND 60

Db 216 TCYEGRLSYRGAGTQTSGAPCQW-----TVEATYRNMTKALSGLGHLGHAFCRNPND 271

QY 61 RRRPWCYVQGLKPLVQECMVHDC-----ADGKKPSPP 94

Db 272 DTRPWCYVQGLKPLVQECMVHDC-----ADGKKPSPP 94

RESULT 18

O97507 PRELIMINARY; PRT; 616 AA.

AC O97507;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE FXII

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA Takahashi T., Kihara T.;

RL Porcine liver factor XII.

RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; AB022426; BAA37148.1; -

DR HSSP; P00763; LDPO.

DR MEROPS; S01-211; -

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000742; EGF_2.

DR InterPro; IPR006209; EGF-like.

DR InterPro; IPR000083; Fibrinctn1.

DR InterPro; IPR000562; FN Type II.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00039; fn1; 1.

DR Pfam; PF00040; fn2; 1.

DR Pfam; PF00051; kringle; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00013; ENTPEII.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000995; FN Type II; 1.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00181; EGF; 2.

DR SMART; SM00058; FN1; 1.

DR SMART; SM00059; FN2; 1.

DR SMART; SM00130; KR; 1.

DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS00022; EGF_1; 2.

DR PROSITE; PS01186; EGF_2; 1.

DR PROSITE; PS01253; FIBRONECTIN_1; 1.

DR PROSITE; PS00023; FIBRONECTIN_2; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00021; KRINGLE_2; 1.

DR PROSITE; PS00070; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

DR PROSITE; PS00240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;

KW Serine protease.

SQ SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBBD1A9 CRC64;

Query Match 32.2%; Score 178.5; DB 6; Length 616;

Best Local Similarity 41.2%; Pred. No. 8.5e-13;

Matches 42; Conservative 7; Mismatches 40; Indels 13; Gaps 4;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLCQTY-HAHRSDALQLGLGKHNYPND 61

Db 217 CYSDRGLSYRGAGTQTSGAPCQWAS-----EATYRNMTAEQALNWGLGHAFCRNPND 272

QY 62 RRPWCYVQGLKPLVQECMVHDCAD--GKKP-----SSPPE 95

Db 273 TRPWCYVQGLKPLVQECMVHDCAD--GKKP-----SSPPE 95

RESULT 19

```
Q8AXX3
ID Q8AXX3 PRELIMINARY; PRT; 421 AA.
AC Q8AXX3
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Muscle-specific receptor tyrosine kinase MUSK.
OS Gallus gallus (Chicken).
DE Kremen2.
GN KRM2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22308873; PubMed=12421700;
RA Davidson G., Mac B., Del Barco Harrantes I., Niehrs C.;
RT "Kremen proteins interact with Dickkopf1 to regulate anteroposterior
RT CNS patterning.";
RL Development 129:5597-5596(2002).
DR EMBL; AY150813; AAN64661.1; -.
SQ SEQUENCE 421 AA; 46274 MW; 74D04DB1682CD15F CRC64;

Query Match 31.1%; Score 172.5; DB 13; Length 421;
Best Local Similarity 50.7%; Pred. No. 2.9e-12;
Matches 34; Conservative 4; Mismatches 26; Indels 3; Gaps 2;

QY 3 CYEGNGHFYRGKAS-TDTMGRCPLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPDR 61
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29 CFTVNGRDYGTGVSQAGPSTGLYNNQTT--QHLNNAQSDPGLGLGNHNYCRNPDDAD 86
QY 62 RRPWCYV 68
Db :|||||
87 VQPCYV 93

RESULT 20
Q81Z25
ID Q81Z25 PRELIMINARY; PRT; 615 AA.
AC Q81Z25
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Coagulation factor XII-Mie.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wada H., Nishioka J., Nakatani K., Kasai Y., Abe Y., Nobori T.;
RT "Molecular characterization of coagulation factor XII-Mie.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB095845; BAC23095.1; -.
SQ SEQUENCE 615 AA; 67735 MW; 030508870A0C7EDB CRC64;

Query Match 30.8%; Score 170.5; DB 4; Length 615;
Best Local Similarity 48.5%; Pred. No. 7.7e-12;
Matches 33; Conservative 6; Mismatches 24; Indels 5; Gaps 2;

QY 2 TCYEGNGHFYRGKASDTMGRCPLPWSATVLOQTY-HAHRSDALQLGLGKHNYCRNPDR 60
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 SCYDGRGLSYRGARTLTLSGAPCPWAS----EATYRNVTAEQARNWGLGHAFCRNPDR 271
QY 61 RRPWCYV 68
Db :|||||
272 DIRPWCYV 279

RESULT 21
Q8AXY6
ID Q8AXY6 PRELIMINARY; PRT; 947 AA.
AC Q8AXY6
```

```
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Muscle-specific receptor tyrosine kinase MUSK.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20538710; PubMed=11083926;
RA Ip F.C., Glass D.G., Gies D.R., Cheung J., Lai K.O., Fu A.K.,
RA Yancopoulos G.D., Ip N.Y.;
RT "Cloning and characterization of muscle-specific kinase in chicken.";
RL Mol. Cell. Neurosci. 16:661-673(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Gies D., Glass D.J., Yancopoulos G.D.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY143173; AAN05008.1; -.
KW Receptor; Kinase.
SQ SEQUENCE 947 AA; 105588 MW; E6C3FCC5796CC8BE CRC64;

Query Match 30.1%; Score 167; DB 13; Length 947;
Best Local Similarity 35.9%; Pred. No. 3.3e-11;
Matches 42; Conservative 8; Mismatches 29; Indels 38; Gaps 8;

QY 1 KTCYEGNGHFYRGKASDTMGRCPLPWSATVLOQTYHAHR-----SDALQLGLGK 51
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
463 RCTCYSGNGQFYQGWANVTASGIPCOKWS-----DQAPHLHRTPTQVFPPELSDA----- 510
QY 52 HNYCRNP-DNRRRPWCYVQVGLKPLV--QECMVHDCADG-----KKPSS-----PP 94
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
511 ENYCRNPGCENERPWCYTK---DPSVTWEYCSVSCGDASLSLGTGRKNGETQNLP 564

RESULT 22
Q90Y90
ID Q90Y90 PRELIMINARY; PRT; 452 AA.
AC Q90Y90
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE KREMEN.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21167372; PubMed=11267660;
RA Nakamura T., Aoki S., Kitajima K., Funakoshi H., Takahashi T.,
RA Matsumoto K., Nakamura T.;
RT "Molecular cloning and characterization of Kremen, a novel kringle-
RT containing transmembrane protein.";
RL Biochim. Biophys. Acta 1518:63-72(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 CUB DOMAIN.
DR EMBL; AB070851; BAB64294.1; -.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kring1; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringie.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS01180; CUB; 1.
```

```
DR PROSITE, PS00021; KRINGLE_1; 1.
DR PROSITE, PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
SQ SEQUENCE 452 AA; 50188 MW; ED24BCDIAF4564E2 CRC64;

Query Match 28.5%; Score 158; DB 13; Length 452;
Best Local Similarity 44.1%; Pred. No. 1.7e-10;
Matches 30; Conservative 9; Mismatches 25; Indels 4; Gaps 2;

Qy 3 CYEGNGHYRGKASTDTM--GRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 60
||| ||| : : ||| ||| : : ||| ||| |||
Dd 30 CYTVGADYRGCTQNTSLDGKPKLFWNE--TFQHPYNTLKYPNGEGGLGHEHNYCRNPDG 87
||| ||| : : ||| ||| : : ||| ||| |||

Qy 61 RRRPWCYV 68
||| |||
Dd 88 DVSPWCYI 95
||| |||

RESULT 23
O46506 PRELIMINARY; PRT; 454 AA.
AC O46506;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Apolipoprotein a (Fragment).
GN BABAPOA.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RA Cox L.A., Jett C., Hixson J.E.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice Site Mutation is Associated with Deletion of a Single Exon in a Null Allele.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE EMIL/GENE/DBJ DATABASES.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF029691; AAB37886.1; -.
DR HSSP; P00747.2PK4.
DR MEROPS; S01.999; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Lipoprotein; Protease;
KW Serine protease.
FT NON TER 1
SQ SEQUENCE 454 AA; 50041 MW; 974E30744C187B2F CRC64;

Query Match 28.5%; Score 158; DB 6; Length 454;
Best Local Similarity 39.8%; Pred. No. 1.7e-10;
Matches 35; Conservative 10; Mismatches 31; Indels 12; Gaps 4;

Qy 3 CYEGNGHYRGKASTDTMGRPCLPWNSATVLTQ--QTYHAHRSDALQLGLGKHNYCRNPDN 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Dd 102 CYHGGQGYRGCSFSTTVTGRTCSQSSMTPHQKRTPENHNDGLTM-----NYCRNPD 156
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 61 RRRPWCYVQGLKPLV--QECMVHDCAD 86
||| ||| : : ||| ||| : : ||| ||| |||
```

```
Db 157 DTGPPWCFT---MDPSVRWEYCNLTRCSD 181

RESULT 24
Q9UIR5 PRELIMINARY; PRT; 113 AA.
AC Q9UIR5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21181705; PubMed=11285247;
RA Ogorekova M., Kraft H.G., Ehnholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV types 6 to 10 domain affect Lp(a) plasma concentrations and have different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158663; AAF03680.1; -.
DR EMBL; AF158662; AAF03680.1; JOINED.
DR HSSP; P00747.1PMK.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON TER 1
FT NON TER 113
SQ SEQUENCE 113 AA; 12685 MW; F3D65681D9B5253A CRC64;

Query Match 28.3%; Score 157; DB 4; Length 113;
Best Local Similarity 34.8%; Pred. No. 4.8e-11;
Matches 39; Conservative 13; Mismatches 36; Indels 24; Gaps 6;

Qy 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATV--LQOTYHAHRSDALQLGLGKHNYCRNP 58
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Dd 9 RCYHNGGQSYRGTSFTTVTGRTCSQSSMTPHRQRTPENVPNDGLTM-----NYCRNP 63
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 59 DNRPRWCYVQGLKPLV--QECMVHDCADG-----KKPS-SPPEE 96
||| ||| : : ||| ||| : : ||| ||| |||
Dd 64 DADTGPWCFT---MDPSIRWEYCNLTRCSDTGTGVVAPPTVIQVPSLGPPE 112
||| ||| : : ||| ||| : : ||| ||| |||

RESULT 25
O18783 PRELIMINARY; PRT; 806 AA.
AC O18783;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Plasminogen.
OS Macropus eugenii (Tamar wallaby).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9315;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=98004511; PubMed=9342350;
RA Lown R.M., Schwartz K., Patchy L.;
RT "Convergent evolution of apolipoprotein(a) in primates and hedgehog.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997(1997).
CC -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
```



```

RESULT 30
Q8WNR1
ID Q8WNR1 PRELIMINARY; PRT; 359 AA.
AC Q8WNR1;
DT 01-VAR-2002 (TrenBLrel. 20, Created)
DT 01-VAR-2002 (TrenBLrel. 20, Last sequence update)
DT 01-VAR-2003 (TrenBLrel. 23, Last annotation update)
DE Plasmidogen (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
NCBI_TaxID=9615;
[1]_TaxID=9615;
RN
RP SEQUENCE FROM N.A.
RA Pirie-Shepherd S.R., Coffman K.T., Resnick D., Chan R., Kisker O.,
RA Folkman J., Waters D.J.;
RT "Angiotatin is Detectable in the Urine of Dogs with Spontaneous Bone
RT Cancer.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; AY069985; AAL56519.1; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003966; Prothrombin.
DR Pfam; PF00051; kringle; 4.
DR PRINTS; PRO0018; KRINGLE.
DR PRINTS; PRO1505; PROTHROMBIN.
DR ProDom; PRO00395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR GlycoProtein; Kringle.
KW GLYCOPROTEIN; Kringle.
FT NON_TER 1
FT NON_TER 359
SQ SEQUENCE 359 AA; 41172 MW; 776D35F4A80BDD9E CRC64;

Query Match 27.3%; Score 151.5; DB 6; Length 359;
Best Local Similarity 44.3%; Pred. No. 8e-10;
Matches 31; Conservative 5; Mismatches 21; Indels 13; Gaps 3

QY 3 CYENGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDAL-----QLGLGKHNCRN 57
Db 282 CYHNGQSYRGTSSTTTIGRKQSSSMT-----PHRHXTPHFPEAGL-TWNYCRN 333
QY 58 PDNRPRPCWY 67
Db 334 PDADKSPWCY 343

Search completed: December 3, 2003, 14:43:01
Job time : 12.8119 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:02 ; Search time 15.1223 Seconds
(without alignments)
1007.637 Million cell updates/sec

Title: US-09-880-503-9
Perfect score: 554
Sequence: 1 KTCYEGNGHYRGKASTDTM.....QECVHDCADGKKSPPEE 96

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

- 1: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1980.DAT.*
- 2: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1981.DAT.*
- 3: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1982.DAT.*
- 4: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1983.DAT.*
- 5: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1984.DAT.*
- 6: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1985.DAT.*
- 7: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1986.DAT.*
- 8: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1987.DAT.*
- 9: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1988.DAT.*
- 10: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1989.DAT.*
- 11: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1990.DAT.*
- 12: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1991.DAT.*
- 13: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1992.DAT.*
- 14: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1993.DAT.*
- 15: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1994.DAT.*
- 16: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1995.DAT.*
- 17: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1996.DAT.*
- 18: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1997.DAT.*
- 19: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1998.DAT.*
- 20: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA1999.DAT.*
- 21: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA2000.DAT.*
- 22: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA2001.DAT.*
- 23: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA2002.DAT.*
- 24: /SIDSI/gcgdata/Geneseq/Geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	554	100.0	96	AAE16550	Human uPA kringle
2	554	100.0	143	AAE16549	Human uPA amino te
3	554	100.0	337	AAE16549	Human colon cancer
4	554	100.0	337	ABP41795	Human ovarian anti
5	554	100.0	365	AAE16549	Delta 1-46 urokinase
6	554	100.0	378	AAW13535	Human prourokinase
7	554	100.0	386	AAE16549	Bifunctional uroki
8	554	100.0	389	AAW13536	Human prourokinase
9	554	100.0	390	AAE16549	Bifunctional uroki

10	554	100.0	390	16	AAE16549	Bifunctional uroki
11	554	100.0	392	16	AAE16549	Bifunctional uroki
12	554	100.0	392	16	AAE16549	Bifunctional uroki
13	554	100.0	392	16	AAE16549	Bifunctional uroki
14	554	100.0	392	16	AAE16549	Bifunctional uroki
15	554	100.0	392	16	AAE16549	Bifunctional uroki
16	554	100.0	392	16	AAE16549	Bifunctional uroki
17	554	100.0	392	16	AAE16549	Bifunctional uroki
18	554	100.0	392	16	AAE16549	Bifunctional uroki
19	554	100.0	392	16	AAE16549	Bifunctional uroki
20	554	100.0	392	16	AAE16549	Bifunctional uroki
21	554	100.0	393	16	AAE16549	Bifunctional uroki
22	554	100.0	393	16	AAE16549	Bifunctional uroki
23	554	100.0	393	16	AAE16549	Bifunctional uroki
24	554	100.0	393	16	AAE16549	Bifunctional uroki
25	554	100.0	393	16	AAE16549	Bifunctional uroki
26	554	100.0	393	16	AAE16549	Bifunctional uroki
27	554	100.0	393	16	AAE16549	Bifunctional uroki
28	554	100.0	393	16	AAE16549	Bifunctional uroki
29	554	100.0	393	16	AAE16549	Bifunctional uroki
30	554	100.0	393	16	AAE16549	Bifunctional uroki
31	554	100.0	393	16	AAE16549	Bifunctional uroki
32	554	100.0	393	16	AAE16549	Bifunctional uroki
33	554	100.0	393	16	AAE16549	Bifunctional uroki
34	554	100.0	393	16	AAE16549	Bifunctional uroki
35	554	100.0	401	11	AAW13637	Human prourokinase
36	554	100.0	410	13	AAE16549	Prourokinase mutan
37	554	100.0	411	6	AAE16549	Sequence encoded b
38	554	100.0	411	11	AAW13634	Human native prour
39	554	100.0	411	11	AAE16549	UK-S3 as encoded b
40	554	100.0	411	11	AAE16549	Urokinase precursor
41	554	100.0	411	11	AAE16549	Human pro-urokinas
42	554	100.0	411	11	AAE16549	Human pro-urokinas
43	554	100.0	411	11	AAE16549	Human pro-urokinas
44	554	100.0	411	12	AAE16549	Human pro-urokinase
45	554	100.0	411	12	AAE16549	Pro-urokinase deri

ALIGNMENTS

RESULT 1
AAE16550
ID AAE16550 standard; Protein; 96 AA.

- XX AAE16550;
- XX AC
- XX DT
- XX 09-APR-2002 (first entry)
- XX DE Human uPA kringle and connecting peptide.
- XX DE Human; urokinase-type plasminogen activator; uPA; therapy; hypertension; stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder; microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma; tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing; clotting disorder; uterine contraction disorder; respiratory disease; adult respiratory distress syndrome; male impotence.
- XX OS Homo sapiens.
- XX PN WO200197752-A2.
- XX XX
- XX PD 27-DEC-2001.
- XX XX
- XX PF 13-JUN-2001; 2001WO-US18976.
- XX PR 20-JUN-2000; 2000US-212874P.
- XX PA (UTYPE-) UNIV PENNSYLVANIA.
- XX PI Cines DB, Higazi AA;
- XX DR WPI; 2002-122240/16.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SQ Sequence 337 AA;

Query Match 100.0%; Score 554; DB 22; Length 337;
 Best Local Similarity 100.0%; Pred. No. 9.4e-45;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPDN 60
 DB 74 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPDN 133
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 96
 DB 134 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 169

RESULT 4
 ID ABP41795
 XX ABP41795 standard; Protein; 337 AA.

XX AC ABP41795;
 XX 22-AUG-2002 (first entry)
 XX Human ovarian antigen HVVCB79, SEQ ID NO:2927.
 DE
 XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; chromosome 10q24.
 XX
 OS Homo sapiens.

XX WO20020677-A1.
 XX 03-JAN-2002.
 XX 07-JUN-2001; 2001WO-US18569.
 XX 07-JUN-2000; 2000US-209467P.
 XX (HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;
 XX WPI: 2002-147878/19.
 DR N-PSDB; ABQ54872.
 DR

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX Claim 11; SEQ ID NO 2927; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 337 AA;

Query Match 100.0%; Score 554; DB 23; Length 337;
 Best Local Similarity 100.0%; Pred. No. 9.4e-45;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPDN 60
 DB 74 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPDN 133
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 96
 DB 134 RRRPWCYVQVGLKPLVQECMVHDCADGKKSPPEE 169

RESULT 5
 ID AAR68854
 XX AAR68854 standard; protein; 365 AA.

XX AC AAR68854;
 XX 25-NOV-2003 (updated)
 DT 22-NOV-1995 (first entry)
 XX Delta 1-46 urokinase.
 DE
 XX Human; des-epidermal growth factor homologous plasminogen activator;
 KW uPA; liver membrane; reduced affinity; BGF homologous; thrombosis;
 KW thrombolytic; increased half-life; urokinase.
 XX
 OS Homo sapiens (engineered).
 XX Key Location/Qualifiers
 FH Misc-difference 1
 FT /note= "amino acids 1-46 of wild-type urokinase"
 FT

have been deleted"

FT XX US376547-A.
 XX PD 27-DEC-1994.
 XX PF 29-JAN-1988; 88US-0150267.
 XX PR 30-JAN-1987; 87US-0008795.
 XX PR 29-JAN-1988; 88US-0150267.
 XX PA (AMHP) AMERICAN HOME PROD CORP.
 XX PI Hung PP, Kalyan NK, Lee SL;
 XX PI WPI; 1995-043464/06.
 XX DR New modified plasminogen activator cpds. - having regions removed
 XX PT to reduce affinity for liver membranes and increase circulation
 XX PT half-life.
 XX PS Claim 1; Page ?; 26pp; English.
 XX CC Amino acid residues 1-46 contain the EGF region of human urokinase.
 XX CC Deletion of this region results in a plasminogen activator with
 XX CC reduced affinity for liver cell membranes; the mutant protein is
 XX CC not cleared from the circulation as rapidly as is wild-type tPA.
 XX CC The specification only gives the sequence around the deletion and
 XX CC not the full-length sequence of "delta 1-46 urokinase"; the
 XX CC sequence in AAR6884 has been obtained by amending a previously
 XX CC disclosed wild-type human urokinase sequence (from W09501427)
 XX CC according to the description given in Example 3.
 XX CC (Updated on 25-MAR-2003 to correct PF field.)
 XX SQ Sequence 365 AA;
 Query Match 100.0%; Score 554; DB 16; Length 365;
 Best Local Similarity 100.0%; Pred. No. 1e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFGYRGKASTDTMGRPCLPNSATVLTQTYHAHRSALQLGLGKHNCRPN 60
 DB 2 KTCYEGNGHFGYRGKASTDTMGRPCLPNSATVLTQTYHAHRSALQLGLGKHNCRPN 61
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
 DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 97
 RESULT 6
 ID AAW13635
 AC AAW13635;
 XX 04-JUN-1997 (first entry)
 XX DE Human prourokinase variant lacking entire EGF domain.
 XX KW Human; prourokinase; hPUK; variant; half-life; increase; EGF;
 XX KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.
 XX OS Homo sapiens.
 XX OS Synthetic.
 XX PH Key Location/Qualifiers
 XX FT Region 1..9
 XX FT /note= "residues 1-9 of native hPUK"
 XX FT Region 10..378
 XX FT /note= "residues 43-411 of native hPUK"
 XX FT Misc-difference 144
 XX FT /note= "corresponds to TAC codon"
 XX

PN EP398361-A.
 XX PD 22-NOV-1990.
 XX PF 18-MAY-1990; 90EP-0109472.
 XX PR 22-FEB-1990; 90JP-0042020.
 XX PR 18-MAY-1989; 89JP-0126433.
 XX PR 03-JUL-1986; 86JP-0156936.
 XX PR 18-FEB-1987; 87JP-0036495.
 XX PR 18-MAY-1989; 89JP-0126434.
 XX PA (GREC) GREEN CROSS CORP.
 XX PI Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;
 XX PI Morita M, Tanabe T;
 XX DR WPI; 1990-350146/47.
 XX DR N-PSDB; AAT61672.
 XX PT Human pro-urokinase variants - deficient in loop regions of
 XX PT epidermal growth factor, showing long blood half-life, as
 XX PT fibrinolytic agent
 XX PS Claim 1; Page -; 22pp; English.
 XX CC New variants of human prourokinase (hPUK) comprise a hPUK deficient
 XX CC in (i) at least part of the first loop region of the epidermal growth
 XX CC factor (EGF) domain; (ii) at least part of the first loop and at
 XX CC least part of the second loop; or (iii) at least part of the third
 XX CC loop. The hPUK variants show an increased blood half-life comparable
 XX CC to that of the whole EGF domain-deficient hPUK variant and urokinase
 XX CC while retaining the same properties as those of hPUK. They have
 XX CC potent thrombolytic activity and very little tendency to cause
 XX CC spontaneous bleeding. The present sequence represents a specific variant
 XX CC of hPUK which lacks the entire EGF domain; the sequence does not
 XX CC appear in the specification and has been created using the
 XX CC wild-type hPUK sequence and the junction sequence after
 XX CC deletion, both of which are given (in Fig 1 and in Fig 2(3),
 XX CC respectively).
 XX SQ Sequence 378 AA;
 Query Match 100.0%; Score 554; DB 11; Length 378;
 Best Local Similarity 100.0%; Pred. No. 1.1e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFGYRGKASTDTMGRPCLPNSATVLTQTYHAHRSALQLGLGKHNCRPN 60
 DB 15 KTCYEGNGHFGYRGKASTDTMGRPCLPNSATVLTQTYHAHRSALQLGLGKHNCRPN 74
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 96
 DB 75 RRRPWCYVQVGLKPLVQECMVHDCADGKSPSPPEE 110
 RESULT 7
 ID AAR66266
 AC AAR66266 standard; protein; 386 AA.
 XX AC AAR66266;
 XX DT 25-MAR-2003 (updated)
 XX DT 17-AUG-1995 (first entry)
 XX DE Bifunctional urokinase variant M33.
 XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 XX KW urokinase; variant; mutein.
 XX OS Synthetic.
 XX PH Key Location/Qualifiers


```
ID AAR66245 standard; protein; 390 AA.
XX AC AAR66245;
XX DT 25-MAR-2003 (updated)
XX DT 17-AUG-1995 (first entry)
XX DE Bifunctional urokinase variant M12.
XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX KW urokinase; variant; mutein.
XX OS Synthetic.
XX PH Key Location/Qualifiers
XX FT Region 1..365
XX FT /label= M4
XX FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
XX FT Disulfide-bond 4..85
XX FT Disulfide-bond 25..67
XX FT Disulfide-bond 56..80
XX FT Disulfide-bond 102..233
XX FT Disulfide-bond 143..159
XX FT Disulfide-bond 151..222
XX FT Disulfide-bond 247..316
XX FT Disulfide-bond 279..295
XX FT Disulfide-bond 306..334
XX FT Disulfide-bond 366..371
XX FT Region /label= X1
XX FT Region /label= Y1
XX DE4323754-C1.
XX PN 15-JUL-1993; 93DE-4323754.
XX PD 01-DEC-1994.
XX PF 15-JUL-1993; 93DE-4323754.
XX PR 15-JUL-1993; 93DE-4323754.
XX PA (CHEF ) GRUENENTHAL GMBH.
XX PI Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX PI Wnendt S;
XX DR WPI; 1995-015191/03.
XX CC New bifunctional urokinase derivs and related plasmids - with
XX CC improved fibrinolytic and thrombin inhibiting activities, for
XX CC treating cardiac and cerebral infarct, pulmonary embolism, etc
XX PS Example 1; Page 10 and Fig 1; 34pp; German.
XX CC Bifunctional urokinase derivatives corresponding to the formula
XX CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX CC are specific examples of such derivs. which have both improved
XX CC fibrinolytic and thrombin-inhibiting activities, compared to known
XX CC plasminogen activators or thrombin inhibitors. The proteins are
XX CC useful as thrombolytic agents, e.g. for treatment of arterial
XX CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX CC and pulmonary embolism.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX CC Sequence 390 AA;
XX SQ Query Match 100.0%; Score 554; DB 16; Length 390;
XX SQ Best Local Similarity 100.0%; Pred. NO. 1.1e-44;
XX SQ Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX SQ 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHYCRNPDN 60
XX SQ 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHYCRNPDN 61
```

```
OY 61 RRRPCYVQVGLKPLVQECMVHDCADGKXPSPPPE 96
Db 62 RRRPCYVQVGLKPLVQECMVHDCADGKXPSPPPE 97

RESULT 10
AAR66247
ID AAR66247 standard; protein; 390 AA.
XX AC AAR66247;
XX DT 25-MAR-2003 (updated)
XX DT 17-AUG-1995 (first entry)
XX DE Bifunctional urokinase variant M14.
XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX KW urokinase; variant; mutein.
XX OS Synthetic.
XX PH Key Location/Qualifiers
XX FT Region 1..365
XX FT /label= M4
XX FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
XX FT Disulfide-bond 4..85
XX FT Disulfide-bond 25..67
XX FT Disulfide-bond 56..80
XX FT Disulfide-bond 102..233
XX FT Disulfide-bond 143..159
XX FT Disulfide-bond 151..222
XX FT Disulfide-bond 247..316
XX FT Disulfide-bond 279..295
XX FT Disulfide-bond 306..334
XX FT Disulfide-bond 366..371
XX FT Region /label= X1
XX FT Region /label= Y1
XX DE4323754-C1.
XX PN 01-DEC-1994.
XX PD 15-JUL-1993; 93DE-4323754.
XX PR 15-JUL-1993; 93DE-4323754.
XX PA (CHEF ) GRUENENTHAL GMBH.
XX PI Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX PI Wnendt S;
XX DR WPI; 1995-015191/03.
XX CC New bifunctional urokinase derivs and related plasmids - with
XX CC improved fibrinolytic and thrombin inhibiting activities, for
XX CC treating cardiac and cerebral infarct, pulmonary embolism, etc
XX PS Example 1; Page 10 and Fig 1; 34pp; German.
XX CC Bifunctional urokinase derivatives corresponding to the formula
XX CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX CC are specific examples of such derivs. which have both improved
XX CC fibrinolytic and thrombin-inhibiting activities, compared to known
XX CC plasminogen activators or thrombin inhibitors. The proteins are
XX CC useful as thrombolytic agents, e.g. for treatment of arterial
XX CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX CC and pulmonary embolism.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX CC Sequence 390 AA;
XX SQ Query Match 100.0%; Score 554; DB 16; Length 390;
```

Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGKASTDTMGRCPLPNSATVLOQTYHAHRSDALQGLGKHNCRPN 60
DB 2 KTCYEGNGHFGKASTDTMGRCPLPNSATVLOQTYHAHRSDALQGLGKHNCRPN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 97

RESULT 11
AAR66254
ID AAR66254 standard; protein; 392 AA.
XX AAR66254;
AC AAR66254;
XX 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX Bifunctional urokinase variant M21.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX Synthetic.
OS
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region 372..392
FT /label= Y1
XX DE4323754-Cl.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-4323754.
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF) GRUENENTHAL GMBH.
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wnendt S;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with
FT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial

CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 392 AA;
Query Match 100.0%; Score 554; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGKASTDTMGRCPLPNSATVLOQTYHAHRSDALQGLGKHNCRPN 60
DB 2 KTCYEGNGHFGKASTDTMGRCPLPNSATVLOQTYHAHRSDALQGLGKHNCRPN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKPKSPPEE 97

RESULT 12
AAR66255
ID AAR66255 standard; protein; 392 AA.
XX AAR66255;
AC AAR66255;
XX 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX Bifunctional urokinase variant M22.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX Synthetic.
OS
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region 372..392
FT /label= X1
FT /label= Y1
XX DE4323754-Cl.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-4323754.
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF) GRUENENTHAL GMBH.
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wnendt S;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with
FT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX Example 1; Page 10 and Fig 1; 34pp; German.
PS

XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 392 AA;
SQ Query Match 100.0%; Score 554; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 97
RESULT 13
AAR66256
ID AAR66256 standard; protein; 392 AA.
XX AC AAR66256;
XX 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX Bifunctional urokinase variant M23.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX Synthetic.
XX Key Location/Qualifiers
FH Region 1..365 /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..371 /label= X1
FT Region 372..392 /label= Y1
XX DE4323754-C1.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-4323754.
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF) GRUENENTHAL GMBH.
XX Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Whendt S;
XX

DR WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 392 AA;
SQ Query Match 100.0%; Score 554; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 97
RESULT 14
AAR66257
ID AAR66257 standard; protein; 392 AA.
XX AC AAR66257;
XX 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX Bifunctional urokinase variant M24.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX Synthetic.
XX Key Location/Qualifiers
FH Region 1..365 /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..371 /label= X1
FT Region 372..392 /label= Y1
XX DE4323754-C1.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-4323754.
XX

PR 15-JUL-1993; 93DE-4323754.

XX (CHEF) GRUENTHAL GMBH.

PA Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wndt S;

XX WPI; 1995-015191/03.

XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX Example 1; Page 11 and Fig 1; 34pp; German.XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR6244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 392 AA;

Query Match 100.0%; Score 554; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRNPDN 61

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96

DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 97

RESULT 15

AAR66258

ID AAR66258 standard; protein; 392 AA.

AC AAR66258;

XX 25-MAR-2003 (updated)

DT 17-AUG-1995 (first entry)

XX Bifunctional urokinase variant M25.

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.

XX Synthetic.

XX Key Location/Qualifiers
FT Region 1..365
FT /label= M4FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222

FT Disulfide-bond 247..316

FT Disulfide-bond 279..295

FT Disulfide-bond 306..334

FT Disulfide-bond 366..371

FT Region /label= X1

FT 372..392

FT /label= Y1

XX DE4323754-C1.

XX 01-DEC-1994.

XX 15-JUL-1993; 93DE-4323754.

XX 15-JUL-1993; 93DE-4323754.

XX (CHEF) GRUENTHAL GMBH.

XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wndt S;

XX WPI; 1995-015191/03.

XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX Example 1; Page 11 and Fig 1; 34pp; German.XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR6244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 392 AA;

Query Match 100.0%; Score 554; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRNPDN 61

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96

DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 97

RESULT 16

AAR66259

ID AAR66259 standard; protein; 392 AA.

AC AAR66259;

XX 25-MAR-2003 (updated)

DT 17-AUG-1995 (first entry)

XX Bifunctional urokinase variant M26.

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.

XX Synthetic.

XX Key Location/Qualifiers
FT Region 1..365
FT /label= M4FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222

FT Disulfide-bond 247..316

FT Disulfide-bond 279..295

FT Disulfide-bond 306..334

FT Disulfide-bond 366..371

FT Region /label= X1

FT 372..392

FT /label= Y1

FT /note= "unglycosylated prourokinase (Ser47-Leu411)"

FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Region /label= X1
 FT FT 366..371
 FT Region 372..392
 FT /label= Y1
 XX
 PN DE4323754-Cl.
 XX
 PD 01-DEC-1994.
 XX
 PF 15-JUL-1993; 93DE-4323754.
 XX
 PR 15-JUL-1993; 93DE-4323754.
 XX
 PA (CHEF) GRUENENTHAL GMBH.
 XX
 PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 XX Wnendt S;
 DR WPI; 1995-015191/03.
 XX

XX New bifunctional urokinase derivs and related plasmids - with
 FT improved fibrinolytic and thrombin inhibiting activities, for
 FT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX
 PS Example 1; Page 11 and Fig 1; 34pp; German.

XX Bifunctional urokinase derivatives corresponding to the formula
 CC M4-XI-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 392 AA;

Query Match 100.0%; Score 554; DB 16; Length 392;
 Best Local Similarity 100.0%; Pred. No. 1.1e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
 Db 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
 Qy 61 RRRPWCYQVGLKPLVQECMVHDCADGKXPSPPEE 96
 Db 62 RRRPWCYQVGLKPLVQECMVHDCADGKXPSPPEE 97

RESULT 17
 AAR66260
 ID AAR66260 standard; protein; 392 AA.
 XX
 AC AA:66260;

DT 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)

XX Bifunctional urokinase variant M27.

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; muteln.

OS Synthetic.

XX Key Location/Qualifiers
 FT Region 1..365
 FT /label= M4

FT Disulfide-bond 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Region 366..371
 FT /label= X1
 FT Region 372..392
 FT /label= Y1
 XX
 PN DE4323754-Cl.
 XX
 PD 01-DEC-1994.
 XX
 PF 15-JUL-1993; 93DE-4323754.
 XX
 PR 15-JUL-1993; 93DE-4323754.
 XX
 PA (CHEF) GRUENENTHAL GMBH.
 XX
 PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 XX Wnendt S;
 DR WPI; 1995-015191/03.
 XX

XX New bifunctional urokinase derivs and related plasmids - with
 FT improved fibrinolytic and thrombin inhibiting activities, for
 FT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX
 PS Example 1; Page 11 and Fig 1; 34pp; German.

XX Bifunctional urokinase derivatives corresponding to the formula
 CC M4-XI-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 392 AA;

Query Match 100.0%; Score 554; DB 16; Length 392;
 Best Local Similarity 100.0%; Pred. No. 1.1e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
 Db 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
 Qy 61 RRRPWCYQVGLKPLVQECMVHDCADGKXPSPPEE 96
 Db 62 RRRPWCYQVGLKPLVQECMVHDCADGKXPSPPEE 97

RESULT 18
 AAR66261
 ID AAR66261 standard; protein; 392 AA.
 XX
 AC AAR66261;

DT 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)

XX Bifunctional urokinase variant M28.

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;

KW urokinase; variant; mutein.
 XX Synthetic.
 OS
 XX
 FH Key Location/Qualifiers
 XX Region 1..365
 FT /label= M4
 FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
 FT Disulfide-bond 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Disulfide-bond 366..371
 FT Region /label= X1
 FT Region /label= Y1
 PN DE4323754-C1.
 XX
 XX 01-DEC-1994.
 XX 15-JUL-1993; 93DE-4323754.
 XX 15-JUL-1993; 93DE-4323754.
 XX (CHEF) GRUENTHAL GMBH.
 XX
 PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 PI Wnendt S;
 XX WPI; 1995-015191/03.
 XX
 XX New bifunctional urokinase derivs and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX
 PS Example 1; Page 11 and Fig 1; 34pp; German.
 XX
 CC Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 392 AA;
 Query Match 100.0%; Score 554; DB 16; Length 392;
 Best Local Similarity 100.0%; Pred. No. 1.1e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPDN 60
 DB 2 KTCYEGNGHFYRGKASTDTMGPRCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPDN 61
 QY 61 RRPWCYVQVGLKPLVQECMWVHDCADGKXSSPPEE 96
 DB 62 RRPWCYVQVGLKPLVQECMWVHDCADGKXSSPPEE 97
 RESULT 19
 AAR66263
 ID AAR66263 standard; protein; 392 AA.
 XX
 AC AAR66263;

XX 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX
 DE Bifunctional urokinase variant M30.
 XX
 KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 XX Synthetic.
 OS
 XX
 FH Key Location/Qualifiers
 XX Region 1..365
 FT /label= M4
 FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
 FT Disulfide-bond 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Disulfide-bond 366..371
 FT Region /label= X1
 FT Region /label= Y1
 PN DE4323754-C1.
 XX
 XX 01-DEC-1994.
 XX 15-JUL-1993; 93DE-4323754.
 XX 15-JUL-1993; 93DE-4323754.
 XX (CHEF) GRUENTHAL GMBH.
 XX
 PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 PI Wnendt S;
 XX WPI; 1995-015191/03.
 XX
 XX New bifunctional urokinase derivs and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX
 PS Example 1; Page 11 and Fig 1; 34pp; German.
 XX
 CC Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 392 AA;
 Query Match 100.0%; Score 554; DB 16; Length 392;
 Best Local Similarity 100.0%; Pred. No. 1.1e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPDN 60
 DB 2 KTCYEGNGHFYRGKASTDTMGPRCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPDN 61
 QY 61 RRPWCYVQVGLKPLVQECMWVHDCADGKXSSPPEE 96
 DB 62 RRPWCYVQVGLKPLVQECMWVHDCADGKXSSPPEE 97

1 KTCYEGNGHFRGKASTDTMGPRCLPWNATVLTQTYHAHRSDALQGLGKKNYCRNPON 60
 2 KTCYEGNGHFRGKASTDTMGPRCLPWNATVLTQTYHAHRSDALQGLGKKNYCRNPON 61
 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96
 62 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 97

RESULT 21
 AAR66244
 ID AAR66244 standard; protein; 393 AA.
 AC AAR66244;
 DT 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 DE Bifunctional urokinase variant M11.
 KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 OS Synthetic.
 FH Key
 FT Region
 FT Location/Qualifiers
 FT 1..365
 FT /label= M4
 FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
 FT Disulfide-bond 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Disulfide-bond 366..374
 FT /label= X1
 FT Region
 FT 375..393
 FT /label= Y1
 DE4323754-C1.
 01-DEC-1994.
 15-JUL-1993; 93DE-4323754.
 15-JUL-1993; 93DE-4323754.
 (CHEF) GRUENENTHAL GMBH.
 Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 Wndt S;
 WPI; 1995-015191/03.
 New bifunctional urokinase derivs and related plasmids - with
 improved fibrinolytic and thrombin inhibiting activities, for
 treating cardiac and cerebral infarct, pulmonary embolism, etc
 Example 1; Page 11 and Fig 1; 34pp; German.
 Bifunctional urokinase derivatives corresponding to the formula
 M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 are specific examples of such derivs. which have both improved
 fibrinolytic and thrombin-inhibiting activities, compared to known
 plasminogen activators or thrombin inhibitors. The proteins are
 useful as thrombolytic agents, e.g. for treatment of arterial
 occlusions, deep vein thrombosis, cardiac and cerebral infarction
 and pulmonary embolism.
 (Updated on 25-MAR-2003 to correct PN field.)
 Sequence 392 AA;
 Query Match 100.0%; Score 554; DB 16; Length 392;
 Best Local Similarity 100.0%; Pred. No. 1.1e-44;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX SQ Sequence 393 AA;
Query Match 100.0%; Score 554; DB 16; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQGLGKHNCRPN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQGLGKHNCRPN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 96
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 97
RESULT 22
ID AAR66249 standard; protein; 393 AA.
XX AC AAR66249;
XX DT 25-MAR-2003 (updated)
XX DT 17-AUG-1995 (first entry)
XX DE Bifunctional urokinase variant M16.
XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX KW urokinase; variant; mutein.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Region 1..365 /label= M4
XX FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
XX FT Disulfide-bond 4..85
XX FT Disulfide-bond 25..67
XX FT Disulfide-bond 56..80
XX FT Disulfide-bond 102..233
XX FT Disulfide-bond 143..159
XX FT Disulfide-bond 151..222
XX FT Disulfide-bond 247..316
XX FT Disulfide-bond 279..295
XX FT Disulfide-bond 306..334
XX FT Region 366..372 /label= X1
XX FT Region 373..393 /label= Y1
XX DE4323754-C1.
XX PD 01-DEC-1994.
XX PF 15-JUL-1993; 93DE-4323754.
XX PR 15-JUL-1993; 93DE-4323754.
XX PA (CHEF) GRUENENTHAL GMBH.
XX PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX PI Wnendt S;
XX DR WPI; 1995-015191/03.
XX XX New bifunctional urokinase derivs and related plasmids - with
XX XX improved fibrinolytic and thrombin inhibiting activities, for
XX XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX XX Example 1; Page 10 and Fig 1; 34pp; German.
XX XX Bifunctional urokinase derivatives corresponding to the formula
XX CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266

CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX XX Sequence 393 AA;
SQ Query Match 100.0%; Score 554; DB 16; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQGLGKHNCRPN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQGLGKHNCRPN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 96
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPEE 97
RESULT 23
ID AAR66250 standard; protein; 393 AA.
XX AC AAR66250;
XX DT 25-MAR-2003 (updated)
XX DT 17-AUG-1995 (first entry)
XX DE Bifunctional urokinase variant M17.
XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX KW urokinase; variant; mutein.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Region 1..385 /label= M4
XX FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
XX FT Disulfide-bond 4..85
XX FT Disulfide-bond 25..67
XX FT Disulfide-bond 56..80
XX FT Disulfide-bond 102..233
XX FT Disulfide-bond 143..159
XX FT Disulfide-bond 151..222
XX FT Disulfide-bond 247..316
XX FT Disulfide-bond 279..295
XX FT Disulfide-bond 306..334
XX FT Region 366..372 /label= X1
XX FT Region 373..393 /label= Y1
XX DE4323754-C1.
XX PD 01-DEC-1994.
XX PF 15-JUL-1993; 93DE-4323754.
XX PR 15-JUL-1993; 93DE-4323754.
XX PA (CHEF) GRUENENTHAL GMBH.
XX PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX PI Wnendt S;
XX DR WPI; 1995-015191/03.
XX XX New bifunctional urokinase derivs and related plasmids - with

PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 393 AA;
SQ
Query Match 100.0%; Score 554; DB 16; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFGKASTDTMGPRCLPWN SATVLQQT YHAHRS DALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFGKASTDTMGPRCLPWN SATVLQQT YHAHRS DALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 97
RESULT 24
AAR66251
ID AAR66251 standard; protein; 393 AA.
XX
XX AAR66251;
XX
XX 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M18.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..372
FT Region /label= X1
FT Region 373..393
FT /label= Y1
XX DE4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF) GRUENENTHAL GMBH.
PA

XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wnendt S;
XX
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 393 AA;
SQ
Query Match 100.0%; Score 554; DB 16; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFGKASTDTMGPRCLPWN SATVLQQT YHAHRS DALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFGKASTDTMGPRCLPWN SATVLQQT YHAHRS DALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEE 97
RESULT 25
AAR66252
ID AAR66252 standard; protein; 393 AA.
XX
XX AAR66252;
XX
XX 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M19.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..372
FT Region /label= X1
FT Region 373..393
FT /label= Y1
XX DE4323754-C1.
XX
XX

PD 01-DEC-1994.
XX
PF 15-JUL-1993; 93DE-4323754.
XX
PR 15-JUL-1993; 93DE-4323754.
XX
PA (CHEF) GRUENTHAL GMBH.
XX
PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wndt S;
XX
DR WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 393 AA;
Query Match 100.0%; Score 554; DB 16; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPE 97
RESULT 26
AAR66253
ID AAR66253 standard; protein; 393 AA.
XX
AC AAR66253;
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M20.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4...85
FT Disulfide-bond 25...67
FT Disulfide-bond 56...80
FT Disulfide-bond 102...233
FT Disulfide-bond 143...159
FT Disulfide-bond 151...222
FT Disulfide-bond 247...316
FT Disulfide-bond 275...295
FT Disulfide-bond 306...334

FT Region 366..372
FT /label= X1
FT Region 373..393
FT /label= Y1
XX
PN DE4323754-Cl.
XX
PD 01-DEC-1994.
XX
PF 15-JUL-1993; 93DE-4323754.
XX
PR 15-JUL-1993; 93DE-4323754.
XX
PA (CHEF) GRUENTHAL GMBH.
XX
PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wndt S;
XX
DR WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 393 AA;
Query Match 100.0%; Score 554; DB 16; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPE 96
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGKPSPPPE 97
RESULT 27
AAR99885
ID AAR99885 standard; peptide; 393 AA.
XX
AC AAR99885;
XX
DT 27-JAN-1997 (first entry)
XX
XX M36: fibrinolytic and anticoagulant activity contg. protein.
XX
KW Thrombin; factor Xa; anticoagulant; thrombolytic; plasminogen;
KW urokinase; activator; streptokinase; staphylokinase; APSAC;
KW anisolated plasminogen streptokinase activator complex; hirudin;
KW hirullin; antistatin; pWLT27; pWS1; pSE8; pMW56.
XX
OS Synthetic.
XX
PN EP712934-A2.
XX
PD 22-MAY-1996.
XX
PF 03-NOV-1995; 95EP-0117316.

XX

PS Example 1; Page 21-22; 37pp; German.

XX

CC Example 1 describes the prodn. of plasmids pSE1 and pSE9
contg. a DNA encoding a chimeric protein with fibrinolytic and
thrombin-inhibiting properties.

CC

CC pSE1 encodes the protein given in AAR99597 and pSE9 encodes

CC

CC the protein given in AAR99596.

XX

SQ Sequence 393 AA;

Query Match 100.0%; Score 554; DB 17; Length 393;

Best Local Similarity 100.0%; Pred. No. 1.1e-44;

Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNCRNPDN 60

Db

3 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNCRNPDN 62

Qy

61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96

Db

63 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 98

RESULT 30

AAR47902

ID AAR47902 standard; Protein; 395 AA.

XX

AC AAR47902;

XX

DT 13-JUL-1994 (first entry)

XX

DE Pro-urokinase derivative.

XX

XX Pro-urokinase; half-life; thrombolytic; thrombosis; fibrinolytic;
factor.

XX

OS Homo sapiens.

XX

FN JP05336965-A.

XX

PD 21-DEC-1993.

XX

PF 17-OCT-1991; 91JP-0269615.

XX

PR 17-OCT-1991; 91JP-0269615.

XX

PA (KYOW) KYOWA HAKKO KOGYO KK.

XX

DR WPI; 1994-030907/04.

DR

N-PSDB; AAQ55771.

XX

FT Novel human pro-urokinase derivs. having long half-life - with

PT

high thrombolytic activity, useful for treatment of thrombosis

XX

PS Disclosure; Page 14; 29pp; Japanese.

XX

CC Sequences (AAQ55771-72) are pro-urokinase derivatives. The products
have an inserted sugar moiety having an amino acid substituted,
depleted or inserted variant around the thrombin cleavage site.

CC

CC They also have a long half-life allowing them to be used in the

CC

CC treatment of thrombosis.

XX

SQ Sequence 395 AA;

Qy

1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNCRNPDN 60

Db

32 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNCRNPDN 91

Qy

61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96

Db

92 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 127

Search completed: December 3, 2003, 14:39:16

Job time : 16.1223 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:43:04 ; Search time 9.8558 Seconds
(without alignments)
1811.566 Million cell updates/sec

Title: US-09-880-503-9

Perfect score: 554

Sequence: 1 KTCYEGNGHFYRGKASTDTM.....QECMVHDCADGKKPSPPEE 96

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	554	100.0	96	10	US-09-880-503-9
2	554	100.0	143	10	US-09-880-503-8
3	554	100.0	337	15	US-10-106-698-6266
4	554	100.0	411	10	US-09-880-503-3
5	554	100.0	431	10	US-09-264-468B-1
6	554	100.0	431	12	US-10-301-822-161
7	554	100.0	431	12	US-10-131-985-21
8	554	100.0	431	14	US-10-076-421-2
9	554	100.0	431	15	US-10-171-311-184
10	554	100.0	431	15	US-10-193-656-4
11	552	99.6	431	12	US-10-247-671-149
12	544	98.2	411	12	US-10-407-821-2
13	510	92.1	88	10	US-09-880-503-1
14	510	92.1	135	10	US-09-880-503-4
15	510	92.1	138	10	US-09-584-186-12

16	510	92.1	138	15	US-10-237-667-12	Sequence 12, Appl
17	510	92.1	138	15	US-10-237-708-12	Sequence 12, Appl
18	510	92.1	138	15	US-10-237-866-12	Sequence 12, Appl
19	510	92.1	138	15	US-10-237-871-12	Sequence 12, Appl
20	510	92.1	138	15	US-10-237-624-12	Sequence 12, Appl
21	510	92.1	403	10	US-09-880-503-6	Sequence 6, Appl
22	226	40.8	527	11	US-09-987-457-18	Sequence 18, Appl
23	226	40.8	527	11	US-09-987-455-19	Sequence 19, Appl
24	226	40.8	562	9	US-09-969-271-7	Sequence 7, Appl
25	226	40.8	562	10	US-09-974-298-145	Sequence 145, App
26	226	40.8	562	12	US-10-443-701-4	Sequence 4, Appl
27	226	40.8	562	12	US-10-193-656-8	Sequence 8, Appl
28	218.5	39.4	326	14	US-10-057-951-3	Sequence 3, Appl
29	218.5	39.4	354	11	US-09-987-457-10	Sequence 10, Appl
30	218.5	39.4	354	11	US-09-987-455-11	Sequence 11, Appl
31	218.5	39.4	372	9	US-09-084-491A-3	Sequence 3, Appl
32	218.5	39.4	372	14	US-10-102-704-3	Sequence 3, Appl
33	218.5	39.4	377	11	US-09-987-455-8	Sequence 8, Appl
34	199.5	36.0	560	10	US-09-912-559-3	Sequence 3, Appl
35	199.5	36.0	560	10	US-09-912-559-4	Sequence 4, Appl
36	199.5	36.0	560	15	US-10-172-712-32	Sequence 32, Appl
37	194.5	35.1	555	15	US-10-172-712-28	Sequence 28, Appl
38	179.5	32.4	343	11	US-09-987-457-14	Sequence 14, Appl
39	179.5	32.4	343	11	US-09-987-455-15	Sequence 15, Appl
40	178.5	32.2	339	11	US-09-987-457-12	Sequence 12, Appl
41	178.5	32.2	339	11	US-09-987-455-13	Sequence 13, Appl
42	175.5	31.7	331	11	US-09-987-457-11	Sequence 11, Appl
43	175.5	31.7	331	11	US-09-987-455-12	Sequence 12, Appl
44	173	31.2	83	12	US-09-981-151A-86	Sequence 86, Appl
45	171	30.9	79	12	US-09-981-151A-85	Sequence 85, Appl

ALIGNMENTS

RESULT 1

US-09-880-503-9
; Sequence 9, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-9

Query Match 100.0%; Score 554; DB 10; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.1e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQO TYHAHRS DALQLGLGKHNYCRNPDN	60
Db	1	KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQO TYHAHRS DALQLGLGKHNYCRNPDN	60
QY	61	RRRPMVCYVQVGLKPLVQECMVHDCADGKKPSPPEE	96
Db	61	RRRPMVCYVQVGLKPLVQECMVHDCADGKKPSPPEE	96

RESULT 2

US-09-880-503-8
; Sequence 8, Application US/09880503